University of Alberta

Development of New Liquid Chromatography Techniques for Mass Spectrometric Analysis of Metabolites and Proteins

by

Zhihui Wen

©

A thesis submitted to the Faculty of Graduate Studies and Research in partial fulfillment of the

requirements for the degree of Master of Science

Department of Chemistry

Edmonton, Alberta
Spring 2006

Library and
Archives Canada

Published Heritage
Branch

395 Wellington Street
Ottawa ON K1A 0N4
Canada

Bibliothèque et
Archives Canada

Direction du
Patrimoine de l'édition

395, rue Wellington
Ottawa ON K1A 0N4
Canada

Your file Votre référence
ISBN: 0-494-13908-0
Our file Notre référence
ISBN: 0-494-13908-0

NOTICE:
The author has granted a non-exclusive license allowing Library and Archives Canada to reproduce, publish, archive, preserve, conserve, communicate to the public by telecommunication or on the Internet, loan, distribute and sell theses worldwide, for commercial or non-commercial purposes, in microform, paper, electronic and/or any other formats.

The author retains copyright ownership and moral rights in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

AVIS:
L'auteur a accordé une licence non exclusive permettant à la Bibliothèque et Archives Canada de reproduire, publier, archiver, sauvegarder, conserver, transmettre au public par télécommunication ou par l'Internet, prêter, distribuer et vendre des thèses partout dans le monde, à des fins commerciales ou autres, sur support microforme, papier, électronique et/ou autres formats.

L'auteur conserve la propriété du droit d'auteur et des droits moraux qui protège cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

In compliance with the Canadian Privacy Act some supporting forms may have been removed from this thesis.

While these forms may be included in the document page count, their removal does not represent any loss of content from the thesis.

Conformément à la loi canadienne sur la protection de la vie privée, quelques formulaires secondaires ont été enlevés de cette thèse.

Bien que ces formulaires aient inclus dans la pagination, il n'y aura aucun contenu manquant.

Canada

Dedicated to my wife Ying and my son David

Abstract

Three liquid chromatography (LC) techniques are developed to combine with mass spectrometry (MS) for proteomics and metabolomics applications. The first technique involves the use of a bifunctional LC column with surface chemistry including an ionic group and a long chain hydrophobic group for multidimensional separation. This new single-column multidimensional LC technique is combined with electrospray ionization (ESI) MS for amino acid analysis in human urine samples. In the second technique, bacterial proteins are digested by trypsin and the digested peptides are injected into LC-ESI MS/MS with a separation window of as short as 2.5 min for identification of specific bacterial proteins which leads to rapid bacterial identification. Finally, a chromatofocusing method with a pH range from 3 to 12 has been developed and the applicability of this technique is demonstrated in milk and serum proteome analysis.

Acknowledgements

I would like to thank my supervisor Dr. Liang Li for his encouragement, guidance, financial support, and patience during the completion of this thesis, also for giving me a chance to work in such a well-equipped, world-class laboratory.

I'd like to thank my committee members for reviewing and commenting on my thesis.

All members of the Li group are also gratefully acknowledged, in no particular order: Dr. Ying Zhang, Dr. Chenjie Ji, Mr. Michael Carpenter, Mr. Andy Lo, Mr. Kevin Guo, Mr. Bryce Young, Mr. Mulu Gebre, Ms. Helen Wang, and Ms. Nan Guo.

Finally, thanks also go to my family of their support during my stay in Edmonton.

Table of Contents

Chapter 1 1

Introduction to Multidimensional Liquid Chromatography and Mass Spectrometry........ 1

1.1. Multidimensional Separation Techniques for Small Molecules 1

1.2. Multidimensional Protein Separation with Ion Exchange or Chromatofocusing LC for Proteome Analysis 4

1.3. MALDI-TOF MS 11

1.4. ESI-Ion Trap MS/MS 19

1.5. Objectives of This Thesis 29

1.6. A Brief Summary of This Thesis 30

1.7. References 30

Chapter 2 38

Single-Column Multidimensional HPLC Separation and its Application for Amino Acid Analysis 38

2.1. Introduction 38

2.2. Experimental 41

  2.2.1. Chemicals 41

  2.2.2. HPLC and Electrospray Ionization Mass Spectrometry 41

2.3. Results and Discussion 42

  2.3.1. Principal of Operation on the Bifunctional HPLC Column 42

  2.3.2. Effects of Different Acids in the Mobile Phase on Separation 44

2.3.3. Effects of Long Chain Ion-Pair Reagents on Retention 48
2.3.4. Running the Bifunctional Column in Single Separation Mode: Reversed Phase or Ion Exchange 60
2.3.5. Multidimensional Separation Using a Bifunctional HPLC Column and Its Application in Urinanalysis 67
2.4. Conclusions 80
2.5. References 83

Chapter 3 86

Rapid Bacterial Identification by LC-ESI MS/MS Analysis of Tryptic Peptides Using a Monlithic Capillary column 86
3.1. Introduction 86
3.2. Experimental 88
3.2.1. Chemicals and Reagents 88
3.2.2. Bacteria Samples 88
3.2.3. Cell Lysis, Protein Digestion, and Sample Cleanup 89
3.2.4. HPLC and Electrospray Ionization Mass Spectrometry 89
3.2.5. Data Processing 90
3.3. Results and Discussion 91
3.3.1. Fast Gradient 91
3.3.2. MS and MS/MS scan speed 93
3.3.3. Reproducibility 97
3.3.4. Sensitivity 105

3.3.5. Rapid Bacterial Identification by Monolithic LC ESI MS/MS 106

3.4. Conclusions 110

3.5. References 110

Chapter 4 113

The Method Development and the Applications of pH-Range

Extended Chromatofocusing 113

4.1. Introduction 113

4.2. Experimental 114

4.2.1. Chemicals and Regents 114

4.2.2. Sample Treatment 115

4.2.3. Reversed Phase HPLC-ESI/ MS (or MSMS) 115

4.2.4. Microbore LC-MALDI QqTOF MS (or MSMS) for peptides............... 116

4.2.5. MALDI MS for proteins.. 116

4.2.6. Database Search...117

4.3. Results and Discussion 118

4.3.1. Internal and External Chromatofocusing 118

4.3.2. pH Linearity of the Hybrid External Chromatofocusing 122

4.3.3. Mechanism of Chromatofocusing 130

4.3.4. Analysis of Standard Proteins with the pH-range Extended

Chromatofocusing 134

4.3.5. Application of Chromatofocusing on Milk Separation 141

4.3.6. Application of Chromatofocusing on Human Serum Separation 144

4.4. Conclusions 153

4.5. References 155

Chapter 5 157

Conclusions and Future Work..157

List of Tables

Table 1.1. Liquid chromatography techniques based on specific properties of proteins..5

Table 2.1. Suppression of long chain ion-pair reagents on HPLC-ESI/MS signal of 10 amino acids. (-) means MS signal suppression. The raw data are the same as those for Figure 2.3. 60

Table 2.2. Summary of R^2 and the linear range of MS signal intensity-concentration correlations, and LODs of the 24 analytes listed in Figure 2.11. HPLC conditions are shown in Figure 2.8........................78

Table 3.1. A reproducibility breakdown from the replicate experiments using a 10 min PSW and *(2, 4)* instrument set-up.....................................102

Table 3.2. Overall sensitivity using two different PSWs. All experiments were performed on a LCQ Deca ion trap MS. Acetic acid was used as the ion-pairing reagent and the flow rate was set at 2.5 μl/min. There are around 1.0 X 10^9 cells in 1 mg *E. coli*. This number was obtained through counting the cells under microscope and was used to estimate the number of *E. coli* cells in samples for cell lysis...................105

Table 3.3. The number of identified proteins from pure bacterial cultures and bacterial mixtures using a 2.5 min PSW. Flow rate was set at 2.5 μL/min. MS/MS data were searched against NCBI bacterial database using MASCOT..108

Table 4.1. Buffer composition (mM) used in this chapter. HCl and NaOH were used to adjust the pH of Buffer (I) or Buffer (II). The pH of buffer A with 2 M urea at around 3.0 and the pH of buffer B with 2 M urea at around 12 were used as the starting buffer and elution buffer, respectively....................122

Table 4.2. The 6 proteins' theoretical pIs and observed pIs in Figure 4.7....................135

Table 4.3. Influence of the flow rate f on the observed pI. The A-B linear gradient was 30 min. The column was 25 cm × 1.0 mm i.d....................138

Table 4.4. Influence of the gradient g on the observed pI. The flow rate was 0.1 mL/min. The column was 25 cm × 1.0 mm i.d....................138

Table 4.5. Masses of proteins identified in the MALDI MS spectra (S/N > 5) shown in Figure 4.10....................144

Table 4.6. Number of proteins identified by approaches (I) and (II) shown in Figure 4.14....................152

List of Figures

Figure 1.1. Schematic titration curve of proteins. 6

Figure 1.2. Schematic pH gradient formed on a cation exchange column during chromatofocusing. 8

Figure 1.3. Schematic focusing effect during chromatofocusing. 10

Figure 1.4. Schematic diagram of MALDI. 13

Figure 1.5. Schematic diagram of the MALDI-linear TOF. 14

Figure 1.6. Schematic diagram of the MALDI-linear TOF with time-lag focusing. 15

Figure 1.7. Schematic diagram of the MALDI-linear TOF with time-lag focusing and reflectron. 18

Figure 1.8. Schematic diagram of ESI. 21

Figure 1.9. The principal diagram of QIT. 22

Figure 1.10. The stability diagram in (a_Z, q_Z) space for a region of simultaneous stability in both r and Z direction. 26

Figure 1.11. Schematic diagram of LTQ mass analyzer. 29

Figure 2.1. Schematic representation of the Primsep 100 bifunctional stationary phase. A^- represents the negatively charged functional group for cation exchange separation mechanism. The long alkyl chain on the right side of A^- is the group for reversed-phase separation mechanism. Details of the chemical structures are not available from the manufacturer. 44

Figure 2.2. The $M+H^+$ base peaks of the mixed-mode HPLC/MS when(A) 5% acetic acid with 0.1% TFA in a 70:30 mixture of water-acetonitrile and (B) 0.3% TFA in a 70:30 mixture of water-acetonitrile was used as buffer B. A Primsep 15cm × 2.1mm i.d. column was run in a mixed-mode HPLC condition with a 30 min linear gradient. Other HPLC conditions for (A) and (B) were the same. The injection amount for each amino acid (1) H and (2) F was 2.5 nmol and 1.0 nmol respectively. The fact that F could not be eluted in (A) was verified through its $M+H^+$ base peak. 47

Figure 2.3. Retention time shift when (I) 5 mM HFBA; (II) 10 mM HFBA; (III) 1 mM TFHA; (IV) 5 mM TFHA; (V) 10 mM TFHA; and (VI) 1 mM HFNA was added to buffer A: 0.05% TFA in a 90:10 mixture of water-acetonitrile, and buffer B: 0.6%TFA in a 80:20 mixture of water-acetonitrile. The retention time shifts are calculated in relation to those determined under the same LC/MS conditions but without the use of any long chain ion-pair reagents. A Primsep 25 cm × 2.1 mm i.d. column was run in a mixed-mode condition with a 45 min linear gradient followed by 15 min 100% B. (X), (Y) and (Z) are base peaks chromatogram when no any long chain ion-pair reagents, 5 mM HFBA, and 5 mM TFHA was added respectively. 1 to 10 in (X), (Y) and (Z) stands for amino acids: C, E, A, V, M, I, K, H, F and R. The injection amount for each amino acid was 0.5 nmol. 53

Figure 2.4. The correlation between pI of amino acids: A, V, C, I, K, Q, M, and R and their retention time shift with (A) 10 mM TFHA and (B) 1 mM HFNA. The raw data are the same as those for Figure 2.3. 56

Figure 2.5. The correlation between retention time shifts when (A) 5 mM TFHA or 10 mM TFHA, respectively, applied; (B) 5 mM TFHA or 1 mM HFNA, respectively, was applied; (C) 10 mM TFHA or 1 mM HFNA, respectively, was applied. The raw data are the same as those for Figure 2.3. 58

Figure 2.6. The $M+H^+$ base peaks of: (A) mixed-mode HPLC/MS and (B) reversed phase HPLC/MS used to separate amino acids: (1) F, (2) H, and (3) K. A Primsep 100 25cm × 2.1 mm i.d. column was used in both modes. A 30 min A-B linear gradient from 0% B to 100% B followed by 15 min at 100%B was applied. The buffers for the mixed-mode HPLC conditions were: 0.05% TFA (buffer A) and 0.5% (buffer B), respectively, in a 85:15 mixture of water-acetonitrile. The buffers for the reversed-phase HPLC conditions were: 0.3% TFA in 95:5 mixture of water-acetonitrile (buffer A) and 0.3% TFA in 20:80 mixture of water-acetonitrile (buffer B). Injections contained 2.5 nmol of each amino acid. 63

Figure 2.7. The $M+H^+$ base peaks of: (A) mixed-mode HPLC/MS and (B) cation exchange HPLC/MS used to separate amino acids (1) C and (2) N on a Primsep 100 25 cm × 2.1 mm i.d. mixed-mode HPLC column. A 30 min A-B linear gradient from 0% B to 100% B followed by 15 min at 100% B was applied in both modes. The buffers for the cation exchange HPLC conditions were: 0.05% TFA in 95:5 mixture of water-acetonitrile (buffer A) and 0.3% (buffer B) respectively in a 95:5 mixture of water-acetonitrile. The buffers for the mixed-mode HPLC conditions are: 0.05% TFA (buffer A) and 0.5% (buffer B) respectively in a 80:20 mixture of water-acetonitrile. Injections contained 2.5 nmol of each amino acid. 66

Figure 2.8. The schematic of the analytical setup for section 2.3.5. A 25 cm × 2.1 mm i.d. Primsep 100 was used in different modes. A 25 min A-B linear gradient from 0% B to 100% B for the cation exchange separation followed by a 15 min B-C linear gradient from 100% B to 100% C for mixed-mode separation and then a 40 min C-D linear gradient from 100% C to 100% D for the reversed phase separation. Buffer A was 0.05% TFA in 92:8 mixture of water-acetonitrile. Buffer B was 0.1% TFA in 92:8 mixture of water-acetonitrile. Buffer C was 0.3% TFA in 80:20 mixture of water-acetonitrile. Buffer D was 0.3% TFA in 20:80 mixture of water-acetonitrile. 68

Figure 2.9. The base peak chromatogram of TR (1), amino acids C (2), Y (3), V (4), K (5) and W (6). (A) without and (B) with a 25:75 mixture of acetic acid/2-propanol at a flow rate of 0.067mL/min was used as the post-column TFA-fixer. Other HPLC conditions are shown in Figure 2.8. The injection for analytes 1-6 was 3.2, 8.5, 5.0, 3.4, 12.0, 0.2 nmol respectively. 72

Figure 2.10. The performance of other post-column additives as the TFA-fixer contrasted with the acetic acid/2-propanol mixture. Except for the different post-column additives shown along the x-axis, all other HPLC conditions were the same as those in Figure 2.9. 73

Figure 2.11. The base peak chromatogram of the 24 analytes separated by multidimensional HPLC separation on one single mixed-mode column. The injection amount of TR, D, C, N, S, Q, T, P, E, G, A, V, M, CT, Y, I, L, F, OT, H, K, R, CS and W are: 1.3, 0.6, 2.8, 1.7, 0.5, 0.6, 1.4, 0.4, 1.1, 4.5, 9.7, 1.3, 1.4, 2.1, 0.8, 0.9, 0.9, 0.4, 3.1, 1.6, 1.3, 3.4, 2.2, and 0.05 nmol, respectively. Other HPLC conditions are shown in Figure 2.8. 75

Figure 2.12. The correlation graph for (I) C and (II) D separated in cation exchange mode, and (III) CT and (IV) L separated in mixed-mode, and (V) CS and (VI) K separated in reversed phase mode. The raw data are the same as those for Table 2.2. 77

Figure 2.13. The TIC of a urine sample separated with the multidimensional separation method on one single mixed-mode HPLC column. HPLC conditions are shown in Figure 2.8. 80

Figure 3.1. Base peak intensity chromatograms of capillary monolithic RP HPLC-ESI MS with different gradient times from 10% B to 35% B; (a) 2.5 min, (b) 5 min, (c) 10 min, (d) 15 min, respectively. Flow rate is 2 µL/min. 92

Figure 3.2. (a) The average numbers of matched peptides and identified proteins obtained from 3 replicate experiments with different PSWs and MS instrument settings. (b) The number of matched peptides and identified proteins from each (2,4) setting plotted against PSW. 95

Figure 3.3. Reproducibility analysis from 3 replicate experiments with different PSWs. Flow rate was 2.0 μL/min and 1 μg of the digest of *E. coli* was injected for each experiment. 99

Figure 3.4. Zoomed in reproducibility of one-dimensional capillary monolithic RP HPLC-ESI-MS/MS with a 10-min HPLC gradient and scan event of 4 MS/MS and 2-microscans. The protein IDs in (a), (b), and (c) X-axis are completely identical to each other. 101

Figure 3.5. Distribution of the identified proteins using different numbers of peptides. 104

Figure 3.6. Bacterial identification score (%) from different PSWs. Flow rate was 2.0 μL/min and 1 μg of the digest of *E. coli* was injected for each experiment. The MS/MS spectra were searched against non-redundant bacterial database from NCBI using MASCOT. 107

Figure 3.7. (a) Diagram showing the distribution of the identified proteins. 1 μg of the tryptic digest from *Agrobacterium tumefaciens, Bacillus cereus* and *Lactococcus lactis* (w/w/w, 3:2:1) was injected. (b) Diagram showing the distribution of the identified proteins. 1 μg of the tryptic digest from *Agrobacterium tumefaciens,* and *Lactococcus lactis* (w/w, 10:1) was injected. Flow rate was 2.5 μL/min and a *(2, 4)* instrument set-up was used for all experiments. The MS/MS spectra were searched against non-redundant bacterial database from NCBI using MASCOT. 109

Figure 4.1. The schematic setup for the hybrid external chromatofocusing. The 2.1 mm i.d. strong cation exchange (the polystyrene-divinylbenzene based column with particle size of 10 µm and pore size of 1000 Å) and 1.0 mm i.d. strong cation exchange column (the polystyrene-divinylbenzene based column with particle size of 8 µm and pore size of 1000 Å) were purchased from Biochrom Labs (Terre Haute, IN, USA) and Vydac (Hesperia, CA, USA), respectively. The column was regenerated by flush the column with 20 column volumes of buffer A. 119

Figure 4.2. The titration curve of (I) 0.001 M HCl pre-equilibrated and (II) pre-equilibrated with buffer (B) listed in Table 4.1. 121

Figure 4.3. (I) and (III) are the pH simulation results based on equation 4.17 when α_i= 0.3 was assumed in equation 4.18 when buffer (I) and (II) in Table 4.1 were applied respectively. (II) and (IV) are the experimental pH curves on 25 cm × 1.0 mm i.d. column when buffer (I) and (II) in Table 4.1 were applied respectively. The flow rate was 0.08 mL/min. 127

Figure 4.4. Chromatograms obtained when buffer (I) and (II) in Table 4.1 were used as the running buffers. The other HPLC conditions were the same: a 15 cm × 2.1 mm i.d. strong cation exchange column was used; the flow rate was 0.2 mL/min; the detection wavelength was 280 nm with reference wavelength at 360 nm. 25 μg α-chymotrypsin (1), 30 μg trypsinogen (2) and 10 μg lysozyme (3) were injected in each run. A 30 min A-B linear gradient was used. 129

Figure 4.5. Computational simulation of the pH changes of the elution buffer passing over the column (the dashed lines) and the focusing path of the protein X with pI 8.0 (the saw-tooth wave). The pH of the start buffer is 3.3, and the pH of the elution buffer is 11.7. The focusing effect led to narrower band width of protein X. 132

Figure 4.6. Chromatograms (A) and (B) obtained when buffer (II) in Table 4.1 was used as the running buffer on a 15 cm × 2.1 mm i.d. strong cation exchange column and a 2.5 cm × 2.1 mm i.d. strong cation exchange column, respectively. The other HPLC conditions were the same: the flow rate was 0.2 mL/min; the detection wavelength was 280 nm with reference wavelength at 360 nm. 25 μg β-lactoglobulin A (1) and 30 μg trypsinogen (2) were injected in each run. A 20 min A-B linear gradient was used. 133

Figure 4.7. Chromatogram of 6 standard proteins separated with the extended pH-range chromatofocusing method. Buffer (II) in Table 4.1 was used as the running buffer on a 25 cm × 1.0 mm i.d. strong cation exchange column. The flow rate was 0.08 mL/min; the detection wavelength was 280 nm with reference wavelength at 305 nm. 60 μg ovalbumin (1), 20 μg β-lactoglobulin A (2), 30 μg α-chymotrypsin (3), 25 μg trypsinogen (4), 60 μg cytochrome C (5), and 10 μg lysozyme (6) were injected in each run. A 30 min A-B linear gradient was used. 134

Figure 4.8. Chromatogram of 3 standard proteins separated with (A) strong cation exchange and (B) the extended pH-range chromatofocusing method on the same column: 25 cm × 1.0 mm i.d. strong cation exchange column and at the same flow rate: 0.10 mL/min. Buffer (II) in Table 4.1 was used for the chromatofocusing. 10 mM phosphate buffer at pH 3.0 was used as the running buffer for the cation exchange separation. 0.5 M KCl was added into the buffer B for the cation exchange. The other HPLC conditions were the same: detection wavelength was 280 nm with reference wavelength at 360 nm; 40 µg β-lactoglobulin A (1), 48 µg trypsinogen (2), and 15 µg lysozyme (3) were injected; A 30 min A-B linear gradient was used. 140

Figure 4.9. The chromatogram of 50 µL pretreated milk separated with the extended pH-range chromatofocusing. The column was 15 cm × 2.1 mm i.d. strong cation exchange column. The flow rate was 0.2 mL/min and a 50 min A-B linear gradient was applied. The detection wavelength was the same: 280 nm with reference wavelength at 360 nm. The buffer (II) listed in Table 4.2 was used as the running buffer for the chromatofocusing separation. 10 fractions were collected for MALDI/MS. 142

Figure 4.10. MALDI MS spectra of the 10 fractions plus the MALDI MS spectrum of the original milk sample with no fractionation. 143

Figure 4.11. The chromatogram of 50 μL pretreated human serum separated with the pH-range extended chromatofocusing. The column was a 15 cm × 2.1 mm i.d. strong cation exchange column. The flow rate was 0.2 mL/min and a 60 min A-B linear gradient was applied. The detection wavelength was the same: 280 nm with reference wavelength at 360 nm. The buffer (II) listed in Table 4.2 was used as the running buffer for the chromatofocusing separation. 146

Figure 4.12. Schematic setup for a reproducibility investigation regarding albumin depletion with extended pH-range chromatofocusing. 147

Figure 4.13. The distributions of the intensity ratios of the paired peptides that were d(2)-^{13}C-formaldehyde labeled and 3 d(0)-^{12}C-formaldehyde labeled in the 3 mixtures: 1/3 d(0)-^{12}C-formaldehyde labeled peptides from albumin depletion run #2, run #3, and run #4, respectively mixed with 1/3 d(2)-^{13}C-formaldehyde labeled peptides from albumin depletion run #1. 149

Figure 4.14. Schematic setup to separate and identify the proteins in the non-albumin fractions from the extended pH-range chromatofocusing by LC-ESI MS/MS on an LCQ Advantage. 151

List of Abbreviations

2D-PAGE	Two-dimensional polyacrylamide gel electrophoresis
ACN	Acetonitrile
CF	Chromatofocusing
CID	Collision-induced dissociation
E.coli	Escherichia coli
ESI	Electrospray ionization
HCCA	α-cyano-4-hudroxycinnamic acid
HPLC	High performance liquid chromatography
IEF	Isoelectric focusing
IEX	Ion exchange chromatography
ISD	In-source decay
LC	Liquid chromatography
LTQ	Linear ion trap
m/z	Mass-to-charge ratio
MALDI	Matrix-assisted laser desorption-ionization
MS	Mass spectrometry
MudPIT	Multidimensional Protein Identification Technology
PSD	Post-source decay
QIT	Quadrupole ion trap
SDS-PAGE	Sodium dodecyl sulfate-polyacrylamide gel electrophoresis
TFA	Trifluoroacetic acid
TIC	Total ion chromatogram
TOF	Time-of-flight
UV	Ultraviolet

Chapter 1

Introduction to Multidimensional Liquid Chromatography and Mass Spectrometry

Liquid chromatography (LC) combined with mass spectrometry (MS) has become an important tool for biochemical analysis. LC provides a versatile means of separation of a great variety of molecules based on their interactions with a stationary phase. A number of stationary phases are available for separation of biomolecules such as proteins, peptides and metabolites. Using a multidimensional LC system where more than one separation mechanism is employed, a complex mixture can be separated into relatively simple constituents for MS detection. In this chapter, a brief introduction to several LC and MS techniques related to my thesis work is presented.

1.1. Multidimensional Separation Techniques for Small Molecules

Some basic terminologies regarding separation as follows should be described before talking about multidimensional separation. The efficiency of a chromatographic column is a measure of the capacity of the column to restrain peak dispersion. Retention factor, also called as capacity factor, is used to describe the migration rate of an analyte on a column. Selectivity is the ratio of retention factor of two species separated on the column. Resolution is the separation of two peaks in terms of their average peak width at base. The peak capacity is the maximum number of separated peaks within a time window.

The drive for the development of multidimensional chromatography is because it can provide access to noticeable improvements in separation resolution which is the square root of the sum of the squares of the resolution in all dimensions, and the peak capacity which is equal to the product of peak capacities in all dimensions, assuming the separation techniques used in all dimensions are orthogonal, i.e., the physical and chemical principals that the separation techniques are based on are completely different.[1] In 1944 two-dimensional paper chromatography was developed to analyze protein hydrolyzates.[2] After this innovation, multidimensional separation has been widely used in small molecule separation.

The basic experiment of the multidimensional HPLC mentioned above involves injecting a sample onto the first-dimensional HPLC system, fractionating the eluate and injecting the fractions onto the second-dimensional HPLC system, and repeating such steps on further-dimensional HPLC systems until a satisfactory separation is attained. There are two modes in multidimensional liquid chromatography: off-line[3] and on-line[4]. Both modes have been practiced for many years and have their own advantages and disadvantages. With the off-line multidimensional HPLC, fractions can be easily collected and concentrated. But the off-line multidimensional HPLC is not automated and it is slow and labor-intensive, and liable to sample loss, degradation, and contamination that will affect the reproducibility and quantitation. The on-line multidimensional HPLC is automated and performed in an enclosed environment. Obviously it overcomes those problems introduced in off-line multidimensional HPLC.

However, on-line multidimensional HPLC systems are more expensive because more hardware, such as multiple pumps and on-line traps to concentrate the fractions, are usually introduced into the system. Moreover, the solvents for the on-line multidimensional HPLC must be compatible and the separation speed of the second-dimensional HPLC should be much faster than that of the first-dimensional HPLC so that the on-line systems can be in phase. Generally, an on-line mode is more desirable than an off-line multidimensional HPLC when throughput and reproducibility are heavily weighted, as in the pharmaceutical industry.

On-line multidimensional separations can be sub-divided into two categories: heart-cutting on-line multidimensional separations and comprehensive on-line multidimensional separations. The heart-cutting multidimensional HPLC is achieved by selectively collecting a small number of fractions of interest from the first-dimensional HPLC and subsequently injecting them onto the second-dimensional HPLC. Comprehensive analysis is performed by discretely collecting the first-dimensional HPLC eluate in regular fractions and injecting the entire fraction onto the second-dimensional HPLC. The heart-cutting on-line multidimensional HPLC technique requires prior knowledge of the sample component retention that is normally determined by prior running of standards of the analytes of interest on an uncoupled HPLC. So the comprehensive on-line multidimensional HPLC technique[5-9] is more popular than the heart-cutting on-line multidimensional HPLC technique[10-12] when a wide range of compounds need to be analyzed. For comprehensive on-line multidimensional HPLC, a minimum sampling rate of 3 to 4 per band of the first-dimensional HPLC must be maintained to ensure most of the resolution of the first-dimensional HPLC is retained.[13]

Mixed-mode HPLC columns were introduced to improve the selectivity of the HPLC method.[14] Mixed-mode columns were widely demonstrated to separate analytes that are too close in physical or chemical characteristics to be separated by one dimensional separation, i.e. to provide an optional substitute for multidimensional separation with one dimensional separation. Normally the mixed mode was operated in one-dimensional HPLC.[15-21] Rarely it was coupled with another HPLC column in series to make a multidimensional separation.[22]

1.2. Multidimensional Protein Separation with Ion Exchange or Chromatofocusing LC for Proteome Analysis

Traditionally two-dimensional polyacrylamide gel electrophoresis (2D-PAGE) was used to separate proteins. 2D-PAGE combines two orthogonal separations of isoelectric focusing (IEF) in the first dimension and sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) in the second.[23-32] To improve resolution of complex protein mixtures and run-to-run reproducibility, multidimensional high performance liquid chromatography (HPLC) has been developed. In this way the separation can be interfaced directly with the ion source of a mass spectrometer. In fact, most developmental efforts over the last several years have been focused on alternative approaches to 2D-PAGE, such as “gel-free” proteomics and protein arrays [33].

Different HPLC techniques separate proteins or peptides in terms of the differences in specific properties of proteins or peptides, as shown in Table 1.1. Ion exchange chromatography (IEX) and chromatofocusing (CF) separate proteins according to the differences in their net surface charge. Because the relatively high salt concentration in the mobile phase is not compatible with mass spectrometry (MS), IEX[34-39] and CF[40-45] are normally used as the first dimensional separation in Multidimensional Protein Identification Technology (MudPIT). Another characteristic that makes the two separations good for the first dimensional separation is that they offer higher loading capacity than other separation techniques like reversed phase HPLC.

Table 1.1 Liquid chromatography techniques based on specific properties of proteins.

Property	Technique
Charge	Ion exchange chromatography (IEX); Chromatofocusing (CF)
Size	Size exclusion
Hydrophobicity	Hydrophobic interaction chromatography; Hydrophilic Interaction Chromatography; Reversed phase chromatography;
Biorecognition	Affinity chromatography

The surface charge on which IEX and CF separation techniques are based is a very important characteristic of proteins. A protein is built up of many amino acids containing weak acidic and basic groups, and some post-translational modification groups. All these ionizable groups in the protein can be charged and can contribute to the surface charge. The charge status of the ionizable groups depend on their pKa values, their structure and their chemical microenvironment as well (i.e. proteins are amphoteric). The surface charge of a protein is determined by its overall charge, charge density and the charge distribution. One of the most important chemical microenvironment factors for a protein is the solution's pH. The surface charge changes gradually as the pH changes. Each protein has its own unique net surface charge versus pH relationship, called a titration curve. This curve reveals how the overall net surface charge of the protein changes according to the pH of the surroundings. Figure 1.1 illustrates 3 protein titration curves. The pH where the protein has zero net surface charge is the protein's isoelectric point (pI). In Figure 1.1 protein (I) has the lowest pI and protein (III) has the highest pI. The slope of the pH titration curve, dz/dpH, is different for each protein, but each protein's dz/dpH has a maximum when the pH equals its pI. The net surface charge of each protein is different at a specified pH, but the polarity of the surface charge shows a similar trend: at a pH above its pI, a protein will be negatively charged and will bind to a positively charged medium or anion exchanger; at a pH below its pI, a protein will be positively charged and bind to a negatively charged medium or cation exchanger. The pI of a protein can be theoretically computed if the amino acid sequence is known [46], or be evaluated through isoelectric focusing-electrophoresis [47] or CF [48-50].

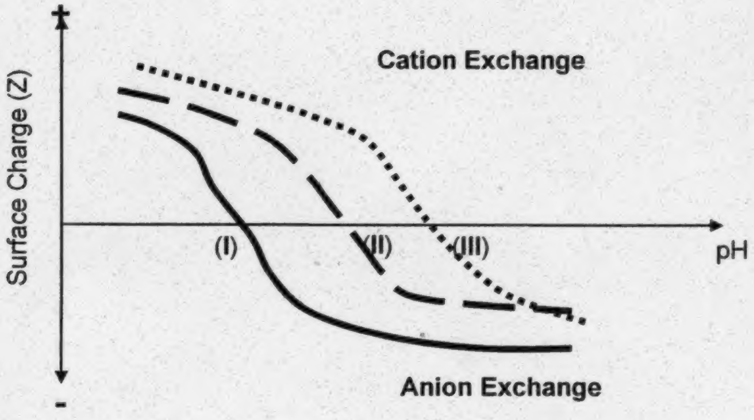


Figure 1.1. Schematic titration curve of proteins.

The most widely used ion exchange separation in proteome analysis is strong cation exchange chromatography. The reason is that the strong ion exchangers maintain their charge over a broader pH range than the weak ion exchangers, allowing selection of the most suitable pH for each application. Two factors determine the resolution of a separation method: selectivity and efficiency. To achieve good selectivity, IEX should be performed at a pH value that the difference in net surface charge of the proteins of interests is maximized. The selectivity of IEX can also be changed by changing the mobile phase composition, column temperature, and composition of the stationary phase. A novel approach to improve the selectivity of IEX, called bipolarity ion-exchange chromatography, is based on combining cation and anion exchangers in the stationary phase.[37] The column efficiency, a measurement of the ability to elute narrow and symmetrical peaks, is another factor that affects the resolution. In general, the smallest particles will produce the narrowest peaks under the correct elution conditions and in a well-packed column, but a smaller particle size often increases the back pressure.

CF is another chromatography method to separate proteins based on the surface charge. The basic idea of CF, introduced about 10 years after the invention of isoelectric focusing by electrophoresis (IEF), is to implement isoelectric focusing with a pH gradient on an IEX column. [51] In order to separate proteins according to their different pIs on a cation exchange column, the column is first equilibrated with a starting buffer at a pH below the lowest pH required so that all proteins can be positively charged and bind to the column. An elution buffer (adjusted to the pH > the highest pI) is passed through the column and begins to titrate the buffer compounds on the column and the proteins. As the buffer flows through the column, the buffer's pH becomes higher while the column's pH becomes lower. Thus a moving and ascending pH gradient shown in Figure 1.2 is generated on the column.

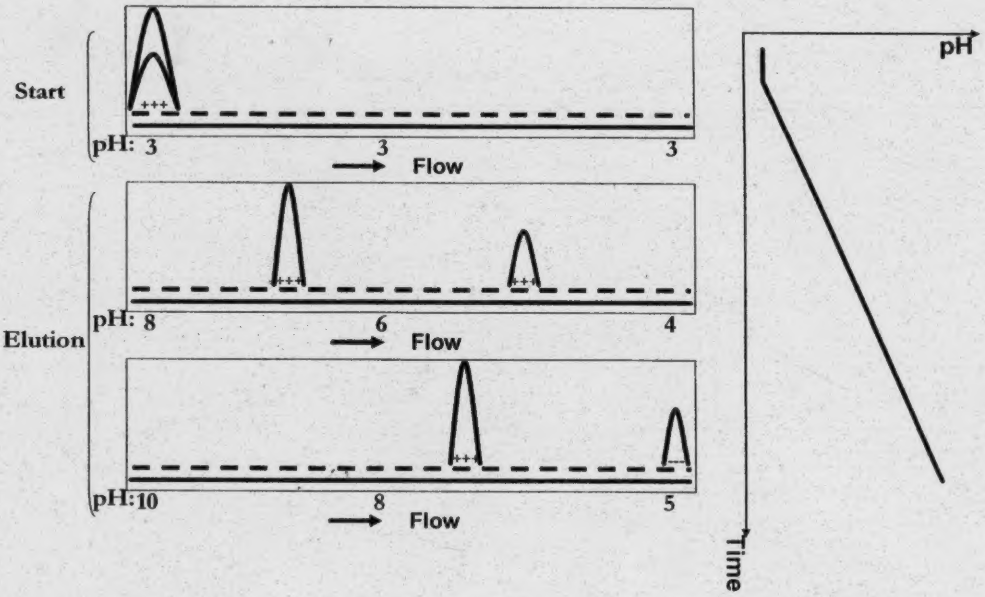


Figure 1.2. Schematic pH gradient formed on a cation exchange column during chromatofocusing.

In Figure 1.2 all proteins initially bind to the cation exchange column. As the pH continues to increase from the top of the column, any protein with a pI below the pH becomes negatively charged and is repelled by the negatively charged column, and begins to migrate down the column with the elution buffer. However, as the protein migrates down the column, the pH of the environs decreases. The protein is titrated by its environs and the negative charge become less and less. When the protein reaches a zone where the pH is below its pI, it becomes positively charged and binds to the column again. The protein remains bound until the developing pH gradient increases the local pH to the point where the protein's pI is less than the local pH again. And then it becomes negatively charged, and begins to move down the column along with the gradient. This titration process continues until the protein is eluted from the column at a pH greater than or equal to its pI, i.e., when it has almost no net charge or a slightly negative charge.

Figure 1.3 illustrates the focusing effect which takes place during chromatofocusing and contributes significantly to the high resolution achievable with CF. In an ascending pH gradient, a protein is constantly and cyclically changing its net surface charge from positive charge, zero charge and to negative charge as the pH gradient develops and the protein travels through different pH zones on the column. Molecules at the rear of the protein band will migrate faster than those at the front because the molecules at the rear are more negatively charged and more repulsed from the column. Gradually in this way a narrower protein band is formed along the column. Thus, during chromatofocusing proteins migrate down the column at different rates that are determined by their pIs as the pH gradient develops (separation selectivity), continually binding to and dissociating from the column and being focused into narrow bands (separation efficiency) and finally eluted. The protein with the lowest pI elutes first and the protein with the highest pI will elute last.

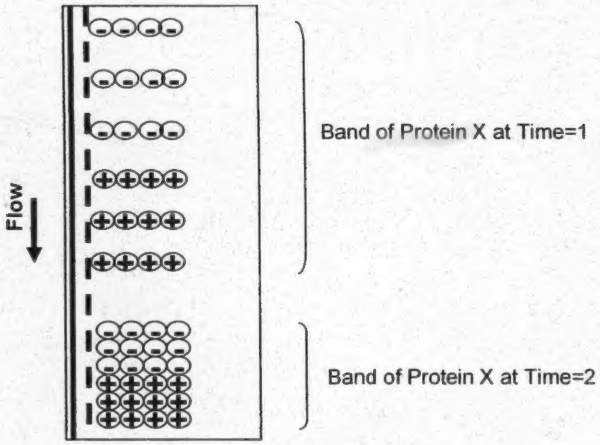


Figure 1.3. Schematic focusing effect during chromatofocusing.

Compared to the gel-based IEF, CF has some obvious advantages: first, much larger sample quantities can be loaded; second, fraction handling such as the subsequent concentration and buffer exchange can be easily accomplished with centrifugal filter devices; and most importantly, the run-to-run reproducibility of CF is much better.[52-54] Compared to IEX, CF also has some obvious advantages: the chromatofocusing method's separation efficiency is higher than IEX due to the focusing effect; because a buffer with low ionic strength is applied, the regeneration time required for the column is far shorter than that for IEX; and the proteins would not be denatured during separation.

The most popular chromatofocusing method is based on the Polybuffer and Polybuffer exchangers from Amersham Pharmacia (Piscataway, NJ, USA).[55-58] A disadvantage of Polybuffer for preparative application is the difficulty to remove the polyampholytes from the eluted protein fractions.[59] In order to achieve consistently high resolution, generation of good pH linearity requires a well-matched buffering capacity over the entire pH range used for a separation. However, the commercial chromatofocusing method is limited to linear coverage of 3-6 pH units, which is inadequate to cover all proteins' pIs in one proteome analysis run. Proteins with pI out of this range will not be retained on or eluted from the column. In the latter case a salt gradient has to be applied after chromatofocusing.[60] However, proteins eluted by the salt gradient usually show poor resolution.

1.3. MALDI-TOF MS

According to current international guidelines (ICH), impurities and degradation products of pharmaceutical drug substances that exceed the threshold of 0.1% must be identified and qualified by appropriate toxicological studies. If the presence of a highly toxic impurity is considered possible, a challenge is the search for impurities that have not or could not have been detected. All separation and detection methods have their limitations, but recently developed HPLC–MS techniques have brought the analyst closer to the point at which it can be reasonably stated that no significant impurity has escaped attention. From this section two types of mass spectrometry (MS) techniques, MALDI-TOF and ESI-Ion trap MS that have been used in this work are introduced.

MS encompasses a family of methods used to obtain accurate masses of ions in the gas phase. MS is one of the most powerful tools for analyzing a broad spectrum of chemical and biological materials. The three most basic procedures for MS are: sample introduction, vaporization and ionization; ion mass analysis; and ion detection. Matrix-assisted laser desorption-ionization (MALDI) is a method to vaporize and ionize species to the gas phase from analytes that are present in a solid or solvent matrix by irradiating the analyte/matrix mixture with a laser beam of a wavelength that is absorbed by the matrix. MALDI is used for introducing and ionizing large molecules, such as polymers[61], peptides and proteins[62-64], into the mass analyzer. The most common mass analyzer for MALDI is the time-of-flight (TOF) analyzer because it has a wide mass range that enables the detection of molecular masses approaching one million Da, that are masses far greater than that of the matrix. TOF also has the feature of high ion throughput, so the ion suppression by the matrix is minimized. Other features of TOF, such as its excellent sensitivity and resolution, make TOF a suitable mass analyzer for complex biological sample even with minimal sample preparation.[65-68]

The mechanism of MALDI is not completely understood, but it is widely accepted that the ionization of protein molecules occurs via a gas phase proton-transfer reaction between the proteins and the protonated matrix molecules that are first evaporated and excited by the focused laser beam, shown in Figure 1.4.[69]

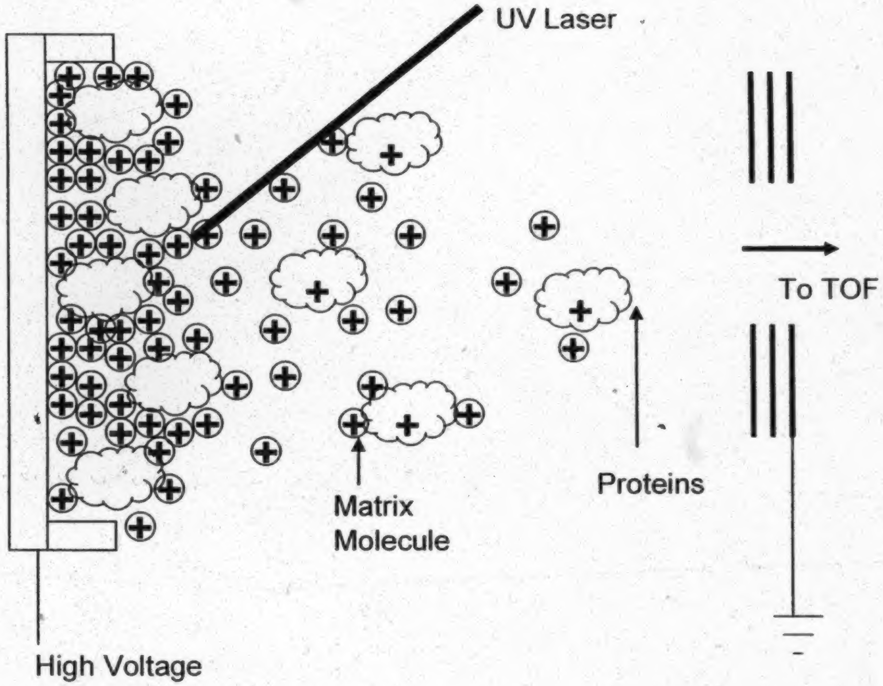


Figure 1.4. Schematic diagram of MALDI.

The challenge encountered in MALDI sample preparation is the so called hot-spot effect where a greater ratio of analytes: crystalline matrix leads to a large variation of signal intensities and, thus, to a poor shot-to-shot reproducibility over a sample spot. Improvement of the homogeneity of MALDI samples can be achieved with fast evaporation[70], deposition by means of a microspot delivery system[71] or electrospray devices[72], the addition of a matrix base[73] and other co-matrixes[74], or using liquid matrixes[75].

A linear TOF instrument is the simplest TOF mass analyzer. It is comprised of a field-free tube for a drift region that is coupled with an ion detector as shown in Figure

1.5. The equations 1.1-1.2 are used to describe the most basic principal of a TOF mass analyzer:

$$zeV = \frac{1}{2}m\upsilon^2 \tag{1.1}$$

$$t = \frac{D}{\upsilon} \tag{1.2}$$

, where z is the charge of the analyte ion, e is the unit of elementary charge, V is the high voltage applied in MALDI (see Figure 1.3), m is the mass of the analyte ion, v is the velocity of the protein along the drift tube, t is the drift time of the protein in the drift tube, and D the length of the field free drift tube. From equations 1.1-1.2, we can get:

$$\frac{m}{z} = t^2\frac{2eV}{D^2} \tag{1.3}$$

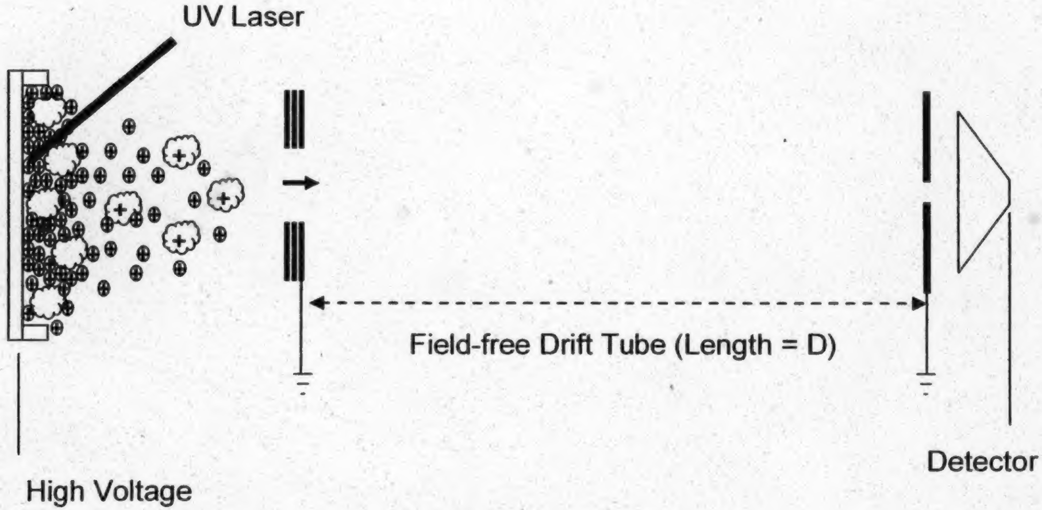

UV Laser
Field-free Drift Tube (Length = D)
Detector
High Voltage

Figure 1.5. Schematic diagram of the MALDI-linear TOF.

The resolution of a linear TOF mass analyzer is very poor due to the broad initial kinetic energy and spatial distribution of protein ions formed in the MALDI process. Time-lag focusing is an energy focusing technique in a TOF mass spectrometer that is accomplished by introducing a time delay between the formation of the ions and the application of the accelerating voltage pulse.[76, 77] Figure 1.6 is the schematic diagram of the MALDI-linear TOF with time-lag focusing.

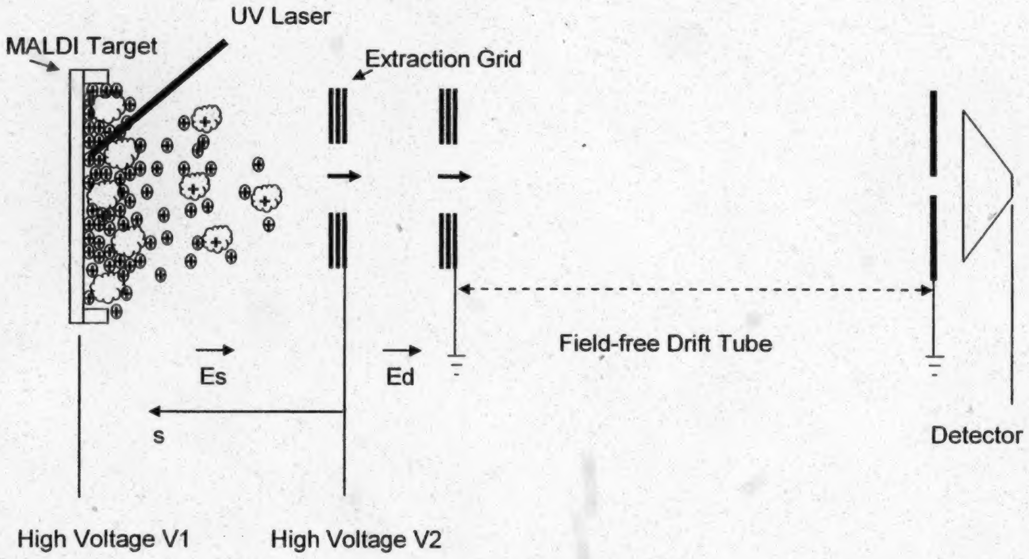


Figure 1.6. Schematic diagram of the MALDI-linear TOF with time-lag focusing.

Consider two identical ions that are produced with different initial energies after the laser pulse is applied. In Figure 1.5 where high voltage on the MALDI target is always applied, they both gain the same amount of energy in the acceleration process. So the ion with higher initial energy will pass the drift tube and reach the detector first. Then the recorded t, i.e. m/z, shows a distribution. Consider now the pulsed case, with an extra grid between the MALDI target and the grounded grid as shown in Figure 1.6. In the desorption-ionization moment, the target and extraction grid are at the same potential. The ion with the higher initial energy will move further out into the space between the target and the grounded grid than the ion with lower initial energy. After a certain delay time, the higher voltage is increased on the target to produce a linear acceleration field between the target and the extraction grid. However, in this case the slower ion will be accelerated more than the faster ion. By choosing correct distances and delay time, the ions will reach the detector at the same time. In this way the recorded distribution of t, i.e. m/z, is corrected. Given the following assumptions: average initial v of each ion equals the same constant; $dt/dv < 0$, where t is the recording time from the target to the detector; and dt/dv is constant; then the delay time τ can be calculated by:

$$\tau = \frac{-1.02 \times \frac{m}{z}}{E_s \times \frac{dt}{ds}} \tag{1.4}$$

where E_s is the electric field strength between the target and the extraction grid, and s is the distance from the extraction grid to the position of the ion when the E_s is applied.

Another better theory was provided to determine the delay time τ.[78] This method was called space velocity correlation focusing because the identification of optimum conditions depends on a correlation between the spatial and velocity distributions of ions at the time of extraction. In brief, the approach involves finding a relationship between the two distributions, using the resulting expression to eliminate a variable v (the initial velocity of the ion) from the equation for t (the recording time from the target to the detector) and then calculating t as a function of the remaining variable. Space velocity correlation focusing theory has a number of advantages over the first time-lag focusing algorithm because it removes the assumptions discussed above. This theory does, however, require other hidden assumptions as the first theory does: it does not include the influence of a metastable ion that is formed with internal energy slightly higher than the threshold for dissociation but with a lifetime long enough to allow it to exit the ion source and travel inside a mass spectrometer dissociating spontaneously before reaching the detector. One must also select a range of presumptive initial velocities (or positions) over which focusing is to be calculated. Finally, the approach requires that a relationship between the initial spatial and velocity distributions that is only valid when spatial distribution of the ions' initial position is ignored, i.e. it has some hidden requirements: all ions are formed in the same location; all ions do not fragment, and all ions do not undergo any collisions that change their velocity. The first hidden requirement can be compensated by presuming a spatial distribution. Regarding the second hidden requirement, decay of ions does take place before or in the field-free drift tube. The first decay is called in-source decay (ISD) and the second decay is called post-source decay (PSD). PSD[79-81], ISD[82-84], and another more routine technique called collision-induced dissociation (CID) that uses inert gas molecules like argon to collide with the target ions, are three basic fragmentation techniques for MALDI-TOF MS/MS. As it is impossible to completely compensate the three hidden requirements, this theory cannot completely correct the distribution of t, i.e. m/z of the same kind of ion.

Significantly higher resolution and improved sensitivity for TOF mass analyzers can be achieved through the so-called reflectron technique.[85] Reflectron is a type of time-of-flight mass spectrometer that uses a static electric field to reverse the direction of travel of the ions entering it. As shown in Figure 1.7, an ion reflectron consists of a series of equally spaced conducting rings that form a retarding field in which the ions penetrate, slow down gradually, and reverse direction, thereby reflecting the ions' trajectory back along the incoming path. The ions with the same mass and charge(s) but with different velocity pass through the drift tube. The faster ions will arrive at the reflectron earlier, but they will penetrate deeper and then take longer to return to the detector,causing ions with the same mass-to-charge ratio but with different velocity to arrive at the detector at the same time. The reflectron decreases the spread in the ion flight times and focuses the ions with the same mass-to-charge ratio, and therefore improves the resolution and sensitivity of the TOF mass analyzer.

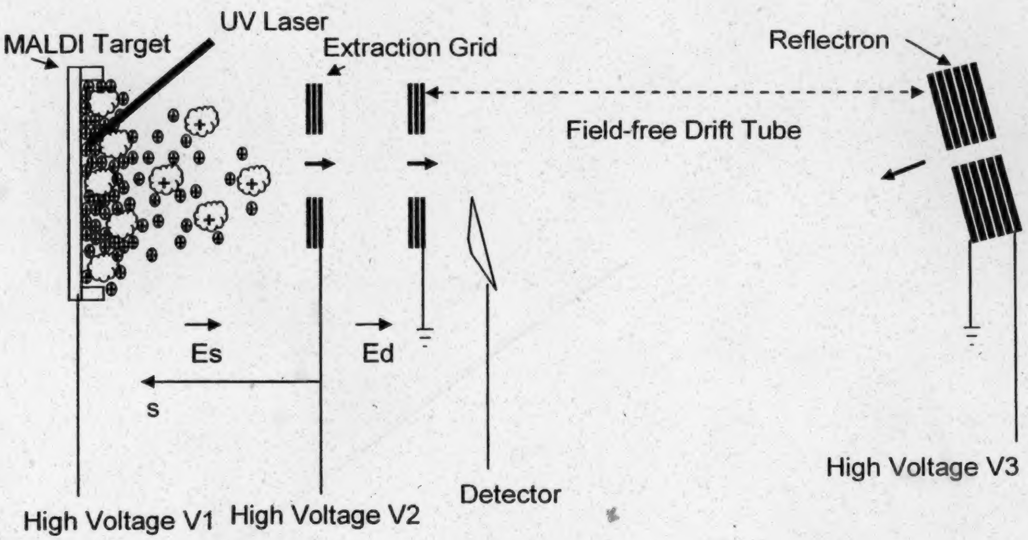

UV Laser
MALDI Target
Extraction Grid
Reflectron
Field-free Drift Tube
Es
Ed
s
High Voltage V3
Detector
High Voltage V1
High Voltage V2

Figure 1.7. Schematic diagram of the MALDI-linear TOF with time-lag focusing and reflectron.

1.4. ESI-Ion Trap MS/MS

Electrospray ionization (ESI) is another method to vaporize and ionize species in the gas phase from analytes that are present in solvent via highly charged fine droplets, by means of spraying the solution from a narrow-bore needle tip at atmospheric pressure in the presence of a high electric field. The pioneering experiments of ESI were made by Malcom Dole et al. who demonstrated the use of electrospray to ionize intact chemical species and thus invent the technique of ESI.[86, 87] Twenty years later John Fenn demonstrated for the first time the use of ESI for the ionization of high mass biologically important compounds and their subsequent analysis by mass spectrometry.[88, 89] This work won John Fenn a share of the 2002 Nobel prize for chemistry. Today ESI has already emerged as an important ionization method for the analysis of proteins, peptides, and small biomolecules with MS.[90-92]

The ESI source has undergone constant development since the earliest examples, but the conceptual arrangement has remained basically the same, as shown in Figure 1.8. The analyte is introduced to the source in solution and passes through an electrospray needle that has a high potential (with respect to the counter electrode) applied (typically in the range from 1.0 to 5.0 kV). The positively charged droplets are repelled from the needle towards the transfer capillary. As the droplets traverse the space, solvent evaporation and ionization of analytes occurs. There is still no complete consensus on the mechanism by which analytes ions are formed from charged droplets. One mechanism is called the charge-residue-model (CRM)[87] and another one is the ion-desorption-model (IDM).[93] The two proposed mechanisms do have some consensus. First, in both models ESI encompasses three successive processes: droplet formation, droplet shrinkage, and gaseous ion formation.[94] At the onset of the electrospray process, the electrostatic force on the liquid causes it to emerge from the tip of the capillary as a jet in the shape of a "Taylor cone". A thin liquid extends from this cone, which breaks into a mist of fine droplets. Several factors such as the applied potential, the flow rate of the solvent, the diameter of the capillary and solvent characteristics influence the size of the initially formed droplets. Evaporation of the solvent from the initially formed droplet as it traverses a pressure gradient toward the analyzer leads to reduction in size, and an increase in surface field, until the Rayleigh limit is reached. A coulomb explosion occurs, as the magnitude of the coulombic repulsion force between the positive charged ions is sufficient to overcome the surface tension holding the shrunken droplets together. The explosion disperses the droplets into a collection of much smaller droplets that continue to evaporate until they too reach the Rayleigh limit and disintegrate again. A simulation of the charged droplet evaporation and fission in ESI confirmed progeny droplets are the primary ion source in ESI.[95] The difference between the two mechanisms is the process of forming ions containing a single analyte molecule. IDM proposes it is the surface electric field that lifts the single analyte ions out, while CRM proposes that it is the solvent evaporation that finally helps form the single analyte ions from the shrunken droplets.

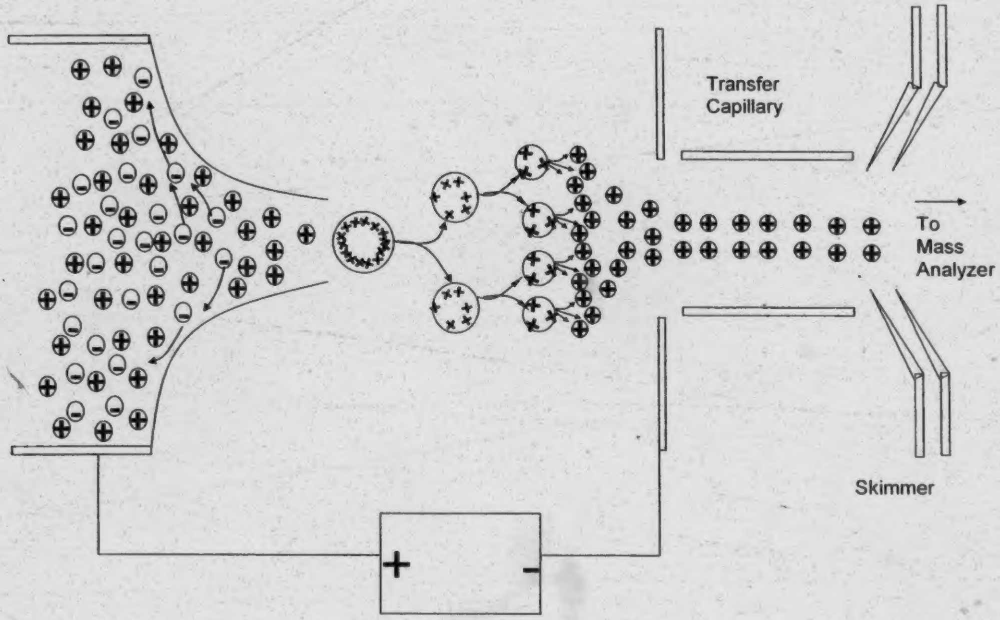


Figure 1.8. Schematic diagram of ESI.

The quadrupole ion trap (QIT) mass analyser was developed in parallel with the quadrupole mass analyser by Wolfgang Paul [96], and it took breakthroughs in manufacturing design in the 1980's [97, 98] to make the QIT-MS/MS the simple and practical instrument that is widely used today.[99-101] The work on QIT won Wolfgang Paul the 1989 Nobel prize for chemistry. The most basic setup of a QIT is shown in Figure 1.9.

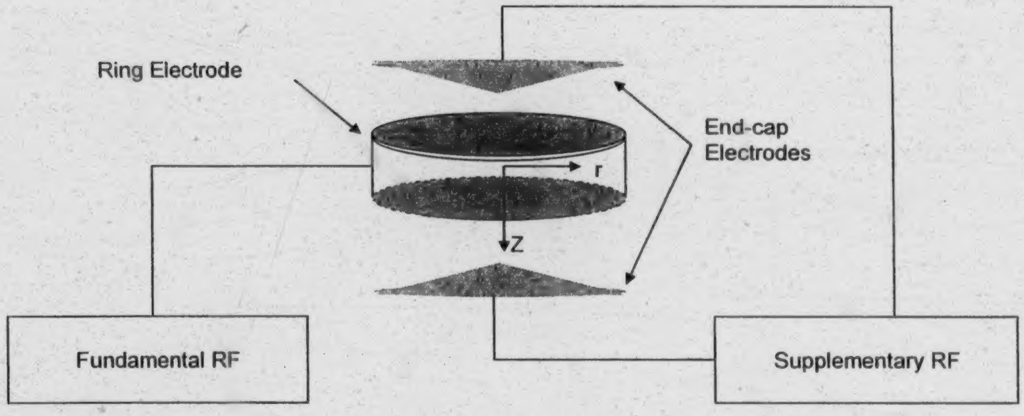

Ring Electrode
End-cap
Electrodes
r
Z
Fundamental RF
Supplementary RF

Figure 1.9. The principal diagram of QIT.

All three electrodes in Figure 1.9 have hyperbolic surfaces that can be defined by the following equations:

End-capelectrodes: $\frac{r^2}{r_0^2} - \frac{Z^2}{Z_0^2} = -1$ (1.4)

Ring electrode: $\frac{r^2}{r_0^2} - \frac{Z^2}{Z_0^2} = 1$ (1.5)

$$r_0^2 = 2Z_0^2 \qquad (1.6)$$

where r_0 is radius of the ring electrode and Z_0 is the distance from the center of the trap to each of the end-cap electrodes.

Suppose the applied RF electric potential to the ring electrode is $\phi_r = U + V\cos(\omega t)$, where $\omega = 2\pi f$ and f is the frequency of the RF, and U and V are the DC and AC component of the fundamental RF in Figure 1.9. When cylindrical coordinates are employed, then the potential at position (r, z) inside the QIT can be expressed as:

$$\phi_{r,Z} = \frac{U+V\cos(\omega t)}{2r_0^2}(r^2 - 2Z^2) \tag{1.5a}$$

For a Quadrupole that has an (x ,y) two-dimensional electric field, Equation 1.5a needs to be expressed as:

$$\phi_{r,Z} = \frac{U+V\cos(\omega t)}{2r_0^2}(r^2 - Z^2) \tag{1.5b}$$

The electric field intensity E_z in z direction, i.e. along the axis through the center of the two end-cap electrodes, can be expressed as:

$$E_Z = -\frac{\partial\phi_{r,Z}}{\partial z} = -\frac{U+V\cos(\omega t)}{r_0^2}(2Z) \tag{1.6}$$

An ion of charge z in the electric field will experience a force F_Z in Z direction that can be expressed as:

$$F_Z = z \times E_Z = z(-\frac{\partial\phi_{r,Z}}{\partial Z}) = z\frac{U+V\cos(\omega t)}{r_0^2}(2Z) \tag{1.7}$$

F_Z is also the product of the ion mass and its acceleration in the Z direction:

$$F_Z = m\frac{d^2Z}{dt^2} \tag{1.8}$$

Inserting Equation 1.7 into Equation 1.8, gives:

$$\frac{d^2Z}{dt^2} = \frac{1}{\frac{m}{z}}\frac{U+V\cos(\omega t)}{r_0^2}(2Z) \tag{1.9}$$

The regions of stability and instability in the QIT in Z direction can be described with the Mathieu equation:[96]

$$\frac{d^2Z}{dt^2} = -\frac{\omega^2}{4}(a_Z - 2q_Z \cos \omega t)Z \tag{1.10}$$

Inserting Equation 1.9 into Equation 1.10 gives:

$$a_Z = -\frac{8U}{\frac{m}{z} r_0^{\ 2} \omega^2} \tag{1.11}$$

$$q_Z = \frac{4V}{\frac{m}{z} r_0^{\ 2} \omega^2} \tag{1.12}$$

When $r_0^{\ 2} \neq 2Z_0^{\ 2}$ in the case of the stretch ion trap, Equation 1.11 and Equation 1.12 can be expressed as:

$$a_Z = -\frac{16U}{\frac{m}{z}(r_0^{\ 2} + 2Z_0^{\ 2})\omega^2} \tag{1.13}$$

$$q_Z = \frac{8V}{\frac{m}{z}(r_0^{\ 2} + 2Z_0^{\ 2})\omega^2} \tag{1.14}$$

In the same way, two similar parameters a_r and q_r to describe the stability and instability in r direction can be obtained where the Mathieu equation is described in Equation 1.15:

$$\frac{d^2{}_r}{dt^2} = -\frac{\omega^2}{4}(a_r - 2q_r \cos \omega t)r \tag{1.15}$$

$$a_r = -\frac{8U}{\frac{m}{z}(r_0^{\ 2} + 2Z_0^{\ 2})\omega^2} \tag{1.16}$$

$$q_r = \frac{4V}{\frac{m}{z}(r_0^{\ 2} + 2Z_0^{\ 2})\omega^2} \tag{1.17}$$

The Mathieu equation 1.10 has two types of solutions: (I) ions oscillate in the Z direction with limited amplitude; (II) the amplitudes in the Z direction grow exponentially. The type (I) solution means ions have stable motion because they will not hit the two cap-end electrodes. The type (II) solution means ions have unstable motion because they will hit the two cap-end electrodes. Similarly, the Mathieu equation 1.15 also has two types of solutions: one for ions in stable motion in r direction and one for ions in unstable motion in r direction. An ion is stable in QIT only when it has stable motion in both Z and r direction. Figure 1.10 is the stability diagram in (a_Z, q_Z) space for a region of simultaneous stability in both r and z directions. When $a_Z=0$, i.e. U in Equation 1.11 equals 0, ions will be selectively scanned out by gradually increasing the RF amplitude in the ring electrode V when $q_Z > 0.908$. Based on Equation 1.12, the m/z selectively scanned out can be expressed as:

$$\frac{m}{z}=\frac{4}{0.908{r_0}^2\omega^2}V \tag{1.18}$$

This scan mode of QIT is called mass-selective instability scan.

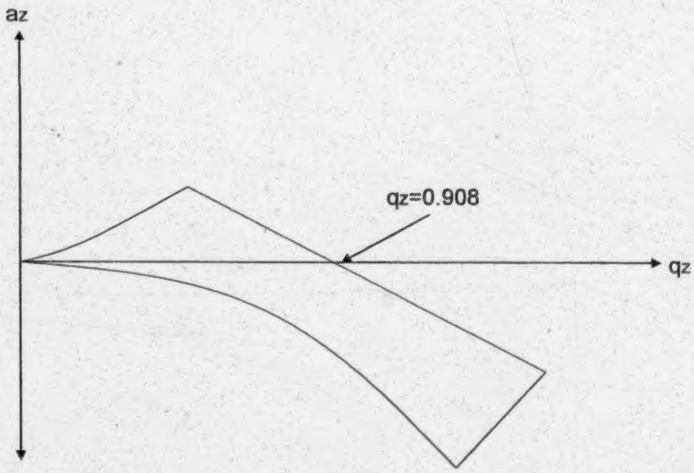


Figure 1.10. The stability diagram in (a_Z, q_Z) space for a region of simultaneous stability in both r and Z direction.

In QIT an ion's trajectory has the general appearance of a " figure 8" composed of two fundamental frequency components in r and z direction: $\omega_{r,n}$ and $\omega_{Z,n}$, which are expressed as:

$$\omega_{r,n} = (n + \frac{1}{2}\beta_r)\omega \tag{1.19}$$

$$\omega_{Z,n} = (n + \frac{1}{2}\beta_Z)\omega \tag{1.20}$$

where n is any non negative integer, and β_r andβ_z are:

$$\beta_r = \sqrt{a_r + \frac{q_r^2}{2}} \tag{1.21}$$

$$\beta_Z = \sqrt{a_Z + \frac{q_Z^2}{2}} \tag{1.22}$$

When a_Z=0 is set, Equation 1.22 and Equation 1.12 are inserted into Equation 1.20 and n=0 is set:

$$\omega_{Z,0} = \frac{1.41V}{\frac{m}{z} r_0^2 \omega} \qquad (1.23)$$

Fixing the parameters (U, V, ω) of the fundamental RF applied on the ring electrode and applied a supplementary RF onto the two cap-end electrodes with frequency f that meets the condition:

$$f = \frac{\omega_{Z,0}}{2\pi} = \frac{0.225V}{r_0^2 \omega} \frac{m}{z} \qquad (1.24)$$

Then the supplementary RF allows the ions to move in resonance with the fundamental RF and pick up some energy in Z direction. If the ions gain enough energy, they will be ejected out of the trap in the Z direction. This is another scan mode of QIT, called resonance ejection. The *f* in Equation 1.24 is termed the secular frequency.

It is possible to selectively isolate a particular m/z in the trap by ejecting all the other ions from the trap. Fragmentation of this isolated precursor ion then can be induced by collision-induced dissociation experiments. The isolation and fragmentation steps can be repeated a number of times and are only limited by the trapping efficiency of the instrument. The very nature of trapping and ejection makes a QIT especially suited to performing MS^n experiments in structural elucidation studies. Space-charge effects (ion-ion repulsion) severely limit the inherent dynamic range of the ion trap, so ESI instead of MALDI is more frequently coupled with QIT.

The original quadrupole mass analyzer is a two-dimensional quadrupole. It lacks MS^n ($n \geq 3$) capability, so the three-dimensional quadrupole mass analyzer QIT has dominated the market for many years. Theoretically a two-dimensional quadrupole mass analyzer reduces the space-charge effects and then can enhance sensitivity.[102] Recently linear ion trap (LTQ) mass analyzers have been developed and commercialized, and are becoming an attractive alternative.[103-105] The LTQ shown in Figure 1.11 is based on the two-dimensional quadrupole mass analyzer. The difference is two additional end-cap electrodes with DC potential applied. LTQ has a very high ion acceptance owing to the absence of a quadrupole field along the z-axis, i.e., the ions are focused to a line rather than to a point in QIT. A comparison of ion capacities of a linear trap vs. a QIT can be expressed as:[106]

$$\frac{N_{LTQ}}{N_{QIT}} = \frac{r_0^{\ 2} l}{Z_0^{\ 3}} \tag{1.25}$$

where N_{LTQ} and N_{QIT} are the trap capacities of LTQ and QIT respectively, l is the length of LTQ. Like QIT, LTQ also has MS^n ($n \geq 3$) capability and resonance ejection scan mode.

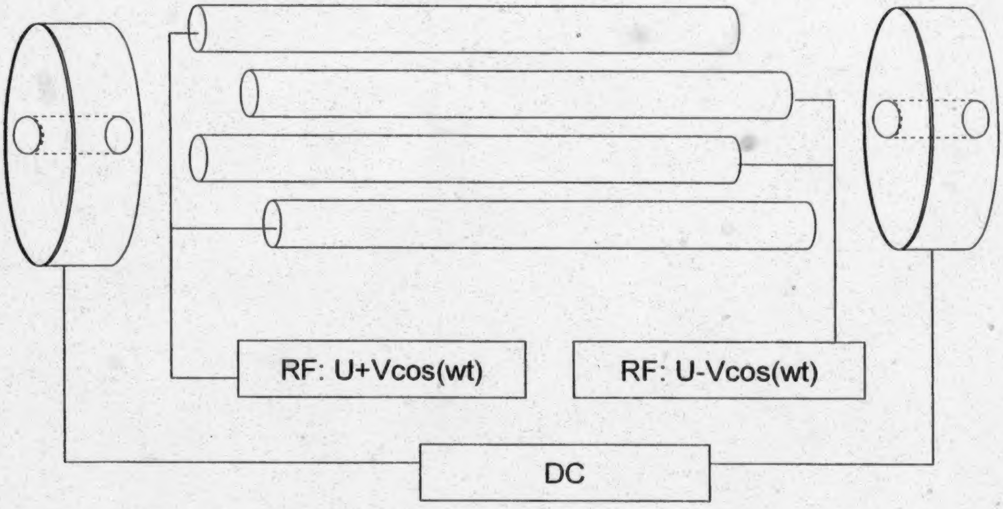


Figure 1.11. Schematic diagram of LTQ mass analyzer.

1.5. Objectives of This Thesis

The general objective of this work is to develop new liquid chromatography techniques for mass spectrometric analysis of metabolites and proteins. Firstly, a MS-friendly multidimensional HPLC method is to be developed for metabolite analysis of urine; secondly, a fast HPLC-ESI MS/MS with high efficiency is to be developed for identification of bacteria; thirdly, a novel HPLC method with high separation efficiency and loading capacity is to be developed for multidimensional HPLC-ESI MS/MS and HPLC-MALDI MS/MS used for identification of protein/peptide.

1.6. A Brief Summary of This Thesis

In Chapter 2 different operation modes of the cation exchange-reversed phase mixed-mode HPLC column were evaluated for separating amino acids. The effects of ion-pair reagents on the separation efficiency of the mixed-mode HPLC column were assessed. A multidimensional separation was achieved with the mixed-mode HPLC column. The method was used to separate 21 amino acids and their metabolites in urine analysis. In Chapter 3 a new method for fast bacterial identification was reported. The fast and efficient separation was achieved through reversed phase HPLC with a capillary monolithic column. In Chapter 4 an extended pH-range CF method and its mechanism were provided and the method was used for milk and serum proteome analysis.

1.7. References

1. Liu Z.; Patterson, D. G.; Lee, M. L. *Anal. Chem.* **1995,** 67, (21), 3840-3845.
2. Consden, R.; Gordon, A. H.; Martin, A. J. P. *Biochem. J.* **1944,** 38, 224-232.
3. Anderson, G. M.; Schlicht, K. R.; Cohen, D. J. *Anal. Chem.* **1983,** 55, (8), 1399-1402.
4. Roston, D. A.; Wijayaratne, R. *Anal. Chem.* **1988,** 60, (9), 948-950.
5. Moore, A. W.; Jorgenson, J. W. *Anal. Chem.* **1995,** 67, (19), 3456-3463.
6. Dugo, P.; Favoino, O.; Luppino, R.; Dugo, G.; Mondello, L. *Anal. Chem.* **2004,** 76, (9), 2525-2530.
7. Holland, L. A.; Jorgenson, J. W. *Anal. Chem.* **1995,** 67, (18), 3275-3283.
8. Graya, M. J.; Dennisa, G. R.; Sloneckerb, P. J; Shalliker, R. A. *J. Chromatogr. A* **2004,** 1041, (1), 101-110.

9. Tanaka, N.; Kimura, H.; Tokuda, D.; Hosoya, K.; Ikegami, T.; Ishizuka, N.; Minakuchi, H.; Nakanishi, K.; Shintani, Y.; Furuno, M.; Cabrera, K. *Anal. Chem.* **2004,** 76, (5), 1273-1281.

10. Posluszny, J. V.; Weinberger, R. *Anal. Chem.* **1988,** 60, (18), 1953-1958.

11. Apffel, J. A.; Alderson, T. V.; Majors, R. E. *J. Chromatogr.* **1981,** 206, (1), 43-57.

12. Sonnefeld, W. J.; Zoller, W. H.; May, W. E.; Wise, S. A. *Anal. Chem.* **1982,** 54, (4), 723-727.

13. Murphy, R. E.; Schure, M. R.; Foley, J. P. *Anal. Chem.* **1998,** 70, (8), 1585-1594.

14. Yamabe, T. *J. Chromatogr.* **1973,** 83, 59-65.

15. Fu, H.; Xie, C.; Xiao, H.; Dong, J; Hu, J; Zou, H. *J. Chromatogr. A* **2004,** 1044, (1), 237-244.

16. Huang, P.; Jin, X.; Chen, Y.; Srinivasan, J. R.; Lubman, D. M. *Anal. Chem.* **1999,** 71, (9), 1786-1791.

17. Lloyd, J. R.; Cotter, M. L; Ohori, D.; Oyler, A. R. *Anal. Chem.* **1987,** 59, (20), 2533-2534.

18. Sun, L; Carr, P. W.; *Anal. Chem.* **1995,** 67, (15), 2517-2523.

19. Ferguson, P. L.; Iden, C. R.; Brownawell, B. J. *J. Chromatogr. A* **2001,** 938, (1), 79-91.

20. Saari-Nordhaus, R.; Anderson J. *Anal. Chem.* **1992,** 64, (19), 2283-2287.

21. Nogueira, R.; Lämmerhofer, M; Lindner, W. *J. Chromatogr. A* **2005,** 1089, (1), 158-169.

22. Venkatramani, C. J.; Zelechonok, Y. *J. Chromatogr. A* **2005,** 1066, (1), 47-53.

23. Chen, Y.; Kim, S. C.; Zhao, Y. *Anal. Chem.* **2005,** 77, (24), 8179-8184.

24. Chromy, B. A.; Gonzales, A. D.; Perkins, J.; Choi, M. W.; Corzett, M. H.; Chang, B. C.; Corzett, C. H.; McCutchen-Maloney, S. L. *J. Proteome Res.* **2004,** 3, (6), 1120-1127.

25. Dougherty, D. A.; Zeece, M. G.; Wehling, R. L. Partridge, J. E. *J. Chromatogr.* **1989,** 480, 359-369.

26. Fernando, A.; Goitom, Z. *J. Chromatogr.* **1991,** 545, (2), 343-347.

27. Klein, J.; Harding, G.; Klein, E. *J. Proteome Res.* **2002,** 1, (1), 41-45.

28. Rammesmayer, G.; Praznik, W. *J. Chromatogr.* **1992,** 623, (2), 399-402.

29. Francis, R. T.; Davie, J.R.; Sayre, M.; Rocha E.; Ziemer, F.; Riedel, G. *J. Chromatogr.* **1984**, 115-121.

30. Thomas, A. H.; and Thomas, J. M. *J. Chromatogr.* **1980,** 195, (2), 297-302.

31. Slater, G. W.; Treurniet, J. R. *J. Chromatogr. A* **1997**, 772, (1), 39-48.

32. Marshall, T; Williams, K. M. *J. Chromatogr. A* **1994,** 662, (1), 167-172.

33. Lambert, J. P.; Ethier, M.; Smith, J. C.; Figeys, D. *Anal. Chem.* **2005,** 77, (12), 3771-3788.

34. Nägele, E.; Vollmer, M.; Hörth, P. *J. Chromatogr. A* **2003,** 1009, (1), 197-205.

35. Gu, S.; Du, Y.; Chen, J.; Liu, Z.; Bradbury, E. M.; Hu, C.-A. A.; Chen, X. *J. Proteome Res.* **2004,** 3, (6), 1191-1120.

36. Millea, K. M.; Kass, I. J.; Cohen, S. A.; Krull, I. S.; Gebler, J. C.; Berger, S. J. *J. Chromatogr. A* **2005,** 1079, (1), 287-298.

37. Ottens, A. K.; Kobeissy, F. H.; Wolper, R. A.; Haskins, W. E.; Hayes, R. L.; Denslow, N. D.; Wang, K. K. *Anal. Chem.* **2005,** 77, (15), 4836-4845.

38. Shen, Y.; Jacobs, J. M.; Camp, D. G.; II; Fang, R.; Moore, R. J.; Smith, R. D.; Xiao, W.; Davis, R. W.; Tompkins, R. G. *Anal. Chem.* **2004,** 76, (4), 1134-1144.

39. Winnik, W. M. *Anal. Chem.* **2005,** 77, (15), 4991-4998.

40. Buchanan, N. S.; Hamler, R. L.; Leopold, P. E.; Miller, F. R.; Lubman, D. M. *Electrophoresis* **2005,** 26, (1), 248-256.

41. Chong, B. E.; Yan, F.; Lubman, D. M.; Miller, F. R. *Rapid Commun. Mass Spectrom.* **2001,** 15, (4), 291-296.

42. Hamler, R. L.; Zhu, K.; Buchanan, N. S.; Kreunin, P.; Kachman, M. T.; Miller, Fred R.; Lubman, D. M. *Proteomics* **2004,** 4, (3), 562-577.

43. Zhu, K.; Kachman, M. T.; Miller, F. R.; Lubman, D. M.; Zand, R. *J. Chromatogr. A* **2004,** 1053, (1), 133-142.

44. Wall, D. B.; Parus, S. J.; Lubman, D. M. *J. Chromatogr. B* **2001,** 763, (1), 139-148.

45. Yan, F.; Subramanian, B.; Nakeff, A.; Barder, T. J.; Parus, S. J.; Lubman, D. M. *Anal. Chem.* **2003,** 75, (10), 2299-2308.

46. May, F. E. B.; Church, S. T.; Major, S.; Westley, B. R. *Biochemistry* **2003,** 42, (27), 8250-8259.

47. Righetti, P. G.; Krishnamoorthy, R.; Gianazza, E.; Labie, D. *J. Chromatogr.* **1978,** 166, (2), 455-460.

48. Gerard, J. P.; Bertoglio J. *J Immunol Methods.* **1982,** 55, (2), 243-251.

49. Mark, A. H.; BODE, K.; COUÉE, I. *Biochem. J.* **1996,** 320, 607-614.

50. Zhu, K.; Zhao, J.; Lubman, D. M.; Miller, F. R.; Barder, T. J. *Anal. Chem.* **2005,** 77, (9), 2745-2755.

51. Elgersma, O.; Sluyterman, L. A. A. E. *J. Chromatogr.* **1978,** 150, (1), 17-30.

52. Romijn, E. P.; Krijgsveld, J.; Heck, A. J. R. *J. Chromatogr. A* **2003,** 1000, (1), 589-608.

53. Shan, L.; Anderson, D. J. *Anal. Chem.* **2002,** 74, (21), 5641-5649.

54. Yan, F.; Subramanian, B.; Nakeff, A.; Barder, T. J.; Parus, S. J.; Lubman, D. M. *Anal. Chem.* **2003,** 75, (10), 2299-2308.

55. Berlin, A. G.; Gusakov, A. V.; Sinitsyna, O. A.; Sinitsyn, A. P. *Appl. Biochem. Biotechnol.* **2000,** 88, (1), 345-352.

56. Deng, F.; Hatzios, K. K. *Pestic. Biochem. Physiol.* **2003,** 74, (2), 102-115.

57. Shimohama, S.; Sasaki, Y.; Fujimoto, S.; Kamiya, S.; Taniguchi, T.; Takenawa, T.; Kimura, J. *Neuroscience* **1998,** 82, (4), 999-1007.

58. Wang, L.; Burhenne, K.; Kristensen, B. K.; Rasmussen, S. K. *Gene* **2004,** 343, (2), 323-335.

59. J. H. Scott.; Kelner, K. L.; Pollard, H. B. *Anal. Biochem.* **1985,** 149, 163-165.

60. Kreunin, P.; Urquidi, V.; Lubman, D. M; Goodison, S. *Proteomics* **2004,** 4, 2754–2765.

61. Guttman, C. M.; Wetzel, S. J.; Flynn, K. M.; Fanconi, B. M.; VanderHart, D. L.; Wallace, W. E. *Anal. Chem.* **2005,** 77, (14), 4539-4548.

62. Li, Li.; Lubman, D. M. *Anal. Chem.* **1988,** 60, (14), 1409-1415.

63. Karas, M.; Hillenkamp, F. *Anal. Chem.* **1988,** 60, (20), 2299-2301.

64. Ohtsu, I.; Nakanisi, T.; Furuta, M.; Ando, E.; Nishimura, O. *J. Proteome Res.* **2005,** 4, (4), 1391-1396.

65. Ochoa, M. L.; Harrington, P. B. *Anal. Chem.* **2005,** 77, (16), 5258-5267.

66. Steele, P. T.; Srivastava, A.; Pitesky, M. E.; Fergenson, D. P.; Tobias, H. J.; Gard, E. E.; Frank, M. *Anal. Chem.* **2005,** 77, (22), 7448-7454.

67. Stemmler, E. A.; Provencher, H. L.; Guiney, M. E.; Gardner, N. P.; Dickinson, P. S. *Anal. Chem.* **2005,** 77, (11), 3594-3606.

68. Wang, Y.; Hornshaw, M.; Alvelius, G.; Bodin, K.; Liu, S.; Sjovall, J.; Griffiths, W. J. *Anal. Chem.* **2005**, DOI: 10.1021/ac051461b.

69. Michael Dale, Richard Knochenmuss, Renato Zenobi, *Rapid Commun. Mass Spectrom.* **1997,** 11, (1), 136-142.

70. Vorm O.; Roepstorff, P.; Mann, M. *Anal. Chem.* **1994,** 66, (19), 3281-3287.

71. Li, L.; Golding, R. E.; Whittal, R. M. *J. Am. Chem. Soc.* **1996,** 118, (46), 11662-11663.

72. Önnerfjord, P.; Ekström, S.; Bergquist, J.; Nilsson, J.; Laurell, T.; Marko-Varga, G. *Rapid Commun. Mass Spectrom.* **1999,** 13, (5), 315-322.

73. Dai, Y.; Whittal, R. M.; Li, L. *Anal. Chem.* **1999,** 71, (5), 1087-1091.

74. Distler, A. M.; Allison, J. *Anal. Chem.* **2001,** 73, (20), 5000-5003.

75. Armstrong, D. W.; Zhang, L.-K.; He, L.; Gross, M. L. *Anal. Chem.* **2001,** 73, (15), 3679-3686.

76. Wiley, W. C.; McLaren I. H. *Rev. Sci. Instrum.* **1955,** 26, (12), 1150-1157.

77. Brown, R. S.; Lennon, J. J. *Anal. Chem.* **1995,** 67, (13), 1998-2003.

78. Colby, S. M.; Reilly, J. P. *Anal. Chem.* **1996,** 68, (8), 1419-1428.

79. Demirev, P. A.; Feldman, A. B.; Kowalski, P.; Lin, J. S. *Anal. Chem.* **2005,** 77, (22), 7455-7461.

80. Mignogna, G.; Giorgi, A.; Stefanelli, P.; Neri, A.; Colotti, G.; Maras, B.; Schinina, M. E. *J. Proteome Res.* **2005,** 4, (4), 1361-1370.

81. Sohn, J.; Parks, J. M.; Buhrman, G.; Brown, P.; Kristjansdottir, K.; Safi, A.; Edelsbrunner, H.; Yang, W.; Rudolph, J. *Biochemistry* **2005,** 44, (50), 16563-16573.

82. Kocher, T.; Engstrom, A.; Zubarev, R. A. *Anal. Chem.* **2005,** 77, (1), 172-177.

83. Neuhof, T.; Schmieder, P.; Preussel, K.; Dieckmann, R.; Pham, H.; Bartl, F.; von Dohren, H. *J. Nat. Prod.* **2005,** 68, (5), 695-700.

84. Suckau, D.; Resemann, A. *Anal. Chem.* **2003,** 75, (21), 5817-5824.

85. Lubman, D. M.; Bell, W. E.; Kronick, M. N. *Anal. Chem.* **1983,** 55, (8), 1437-1440.

86. Mack, L. L.; Kralik, P.; Rheude, A.; Dole, M. *J. Phys. Chem.* **1970,** 52, (10), 4977-4986.

87. Dole, M.; Mack, L. L.; Hines, R. L.; Mobley, R. C.; Ferguson, L. D.; Alice, M. B. *J. Phys. Chem.* **1968,** 49, (5), 2240-2249.

88. Masamichi, Y; Fenn, J. B. *J. Phys. Chem.* **1984,** 88, (20), 4451 - 4459.

89. Fenn, J. B.; Mann, M.; Meng, C. K.; Wong, S. F.; Whitehouse,C. M. *Science* **1989,** 246, 64-71.

90. Fu, Q.; Li, L. *Anal. Chem.* **2005,** 77, (23), 7783-7795.

91. Shen, Y.; Smith, R. D.; Unger, K. K.; Kumar, D.; Lubda, D. *Anal. Chem.* **2005,** 77, (20), 6692-6701.

92. Wang, Y.; Rudnick, P. A.; Evans, E. L.; Li, J.; Zhuang, Z.; DeVoe, D. L.; Lee, C. S.; Balgley, B. M. *Anal. Chem.* **2005,** 77, (20), 6549-6556.

93. Iribarne, J. V.; Thomson,B. A. *J. Chem. Phys.* **1976,** 64, (6), 2287-2294.

94. Tang, L.; Kebarle. P. *Anal. Chem.* **1993,** 65, (24), 3654-3668.

95. Tang, K.; Smith,R. D. *Int. J. Mass Spectrom.* **1999,** 185, 97-105.

96. Paul, W. *Angew. Chem. Int. Ed.* **1990,** 29, (7), 739-748.

97. Stafford, G. C.; Kelley, P. E.; Syka, J. E. P.; Reynolds, W. E.; Todd, J. F. J. *Int. J. Mass Spectrom. Ion Processes* **1984,** 60, (1), 85-98.

98. Louris, J. N.; Cooks, R. G.; Syka, J. E. P.; Kelley, P. E.; Stafford, G. C.; Todd, J. F. J. *Anal. Chem.* **1987,** 59, (13), 1677-1685.

99. Clowers, B. H.; Hill, H. H. Jr. *Anal. Chem.* **2005,** 77, (18), 5877-5885.

100. Crowe, M. C.; Brodbelt, J. S. *Anal. Chem.* **2005,** 77, (17), 5726-5734.

101. Eichhorn, P.; Ferguson, P. L.; Perez, S.; Aga, D. S. *Anal. Chem.* **2005,** 77, (13), 4176-4184.

102. Hager, J. W.; *Rapid Commun. Mass Spectrom.* **2002,** 16, (6), 512-526.

103. Diehnelt, C. W.; Dugan, N. R.; Peterman, S. M.; Budde, W. L. *Anal. Chem.* **2005,** ASAP Article; DOI: 10.1021/ac051556d.

104. Jin, W.H.; Dai, J.; Li, S. J.; Xia, Q. C.; Zou, H. F.; Zeng, R. *J. Proteome Res.* **2005,** 4, (2), 613-619.

105. Yates, J. R.; Cociorva, D.; Liao, L.; Zabrouskov, V. *Anal. Chem.* **2005,** ASAP Article; DOI: 10.1021/ac0514624.

106. Campbell, J. M.; Collings, B. A.; Douglas, D. J. *Rapid Commun. Mass Spectrom.* **1998,** 12, (20), 1463-1474.

Chapter 2

Single-Column Multidimensional HPLC Separation and its Application for Amino Acid Analysis

2.1. Introduction

Multidimensional chromatography has become a powerful separation tool since its introduction in a form of paper chromatography.[1] It provides higher resolution and peak capacity, compared to one-dimensional chromatography.[2] For separating small molecules, the most widely used multidimensional techniques are based on high performance liquid chromatography (HPLC) and gas chromatography (GC).[3-12] One distinct advantage of HPLC over GC for multidimensional separation lies in the great potential of selection and combination of different kind of columns, mobile phases, and additives to provide the needed separation selectivity and peak capacity. In the new era of biosystems analysis where complicated mixtures, such as proteome and metabolome, are studied, multidimensional chromatography is poised to play an increasingly important role. The focus of our research is to develop robust separation techniques that can be readily combined with mass spectrometry (MS) for metabolomics applications.

For the analysis of small biomolecules, such as metabolites, several different separation modes including reversed phase, normal phase, ion exchange, and size exclusion can be combined to form either off-line or on-line multidimensional HPLC.[13]

The idea of using a mixed-mode column for separation has also been explored and it can be traced back to as early as 1973 when cation exchange/anion exchange mixed-mode HPLC was first introduced.[14] Another mixed-mode HPLC, absorbance/reversed phase HPLC, was reported in 1981, where, for efficient separation of peptide and protein mixtures, Hancock et al. used a packing material with low C18-coating and no secondary capping which contained significant concentrations of both free silanol and hydrocarbon groups to allow a mixed-mode separation via adsorption and reversed-phase separation mechanisms.[15] Since then, mixed mode stationary phases have been widely used to separate peptides and small molecules in one-dimensional separations[16-22], but rarely coupled with other HPLC columns for online multidimensional separation[23] where achieving orthogonal separation mechanisms in two or more dimensions is important. In a related work, the use of two or more packed columns with different stationary phases linked in serial, or the use of one tube segmented with different stationary phases, has been demonstrated to be useful for separating complicated mixtures such as proteome digests and small molecules.[19, 24]

A multidimensional HPLC system usually consists of several columns, pumps, valves, and detectors. A multidimensional on-line HPLC system may include one or several intermediate traps to refocus the analytes to correct their sensitivity deterioration caused by analyte dilution along the columns. It would be ideal if multidimensional HPLC could be performed with one column using a HPLC system consisting of only a pair of gradient pumps. First, it would lower the hardware cost significantly by eliminating the need for the multiplex HPLC systems, columns and traps. Second, the band dispersion may be minimized because the analytes do not need to pass over the entire HPLC column before a second-dimensional HPLC is started. When the first-dimensional HPLC is finished, some analytes in the mixture may have been separated well and eluted, others may stay somewhere in the column and wait for the further-dimensional HPLC to separate and elute them. To our knowledge, multidimensional separation using one HPLC column packed with a uniform stationary phase or single-column multidimensional (SCMD) HPLC has not been explored.

SCMD HPLC bears some analogy to earlier work of performing multidimensional separation in one medium. In 1944, Consden, Gordon, and Martin implemented two-dimensional paper chromatography to analyze protein hydrolyzates by applying multiple solvent combinations to one medium. Since then, two-dimensional separation with planar medium chromatography, such as thin-layer chromatography and paper chromatography, has been successfully demonstrated.[25-28] In our work, a commercially available, HPLC column made of bifunctional stationary phase is investigated to examine its suitability for performing different modes of separation in a single column using one set of solvent pumping system. It is demonstrated that this method of separation can be used to separate mixtures of very hydrophobic molecules and very hydrophilic molecules, such as amino acids, without pre-column derivatization or the use of ion-pair additives. The method is compatible with electrospray ionization MS for detection and quantification of amino acids.

2.2. Experimental

2.2.1. Chemicals

20 amino acid standards, carnosine (CS), taurine (TR), ornithine (OT), carnitine (CT), heptafluorobutyric acid (HFBA), tridecafluoroheptanoic acid (TFHA), and heptadecafluorononanoic acid (HFNA) were purchased from Sigma-Aldrich (Oakville, ON, Canada). Trifluoroacetic acid (TFA), acetic acid, propionic acid, acetonitrile and 2-propanol were purchased from Fisher Scientific (Ottawa, ON, Canada). Formic acid was purchased from Anachemia Science (Lachine, QC, Canada). Stock solutions of amino acid standards, CS, TR, OT and CT were prepared in a 95:5 (all ratios are in v:v throughout the chapter) mixture of water-acetonitrile. A urine sample donated by a healthy adult was preliminarily filtered through an Amicon Ultra-15 at 3000 g on a Beckman J2-21 centrifuge (Mississauga, ON, Canada) at 4°C. The filter was purchased from Millipore (Mississauga, ON, Canada). The supernatant was diluted 1:9 with a 95:5 mixture of water-acetonitrile for injection. All stock solution and samples were stored at 4 °C. All water used in this work was purified by a Milli-Q UV plus ultrapure system from Millipore and was 18.3 MΩ cm in specific resistance or better.

2.2.2. HPLC and Electrospray Ionization Mass Spectrometry

The silica based Primsep 100 columns (15 cm × 2.1 mm i.d or 25 cm × 2.1 mm i.d. with particle size of 5 μm and pore size of 100 Å) with bifunctional stationary chemistries for cation-exchange and reversed-phase retention mechanisms, purchased from SIELC Technologies (Prospect Heights, IL, USA), were used for HPLC-ESI MS. The HPLC flow rate was 0.2 mL/min. Ion spectra were acquired on an LCQ Advantage ion trap mass spectrometer from ThermoFinnigan (San Jose, CA, USA) that was coupled with an Agilent (Palo Alto, CA) 1100 HPLC system with a binary pump for gradient generation and solvent delivery. For the work presented in section 2.3.5, the experiments were performed on a Bruker/Agilent Esquire-HPLC Ion Trap HPLC/MS^n system (Silberstreifen, Rheinstetten, Germany) coupled with an Agilent 1100 HPLC system with a quaternary pump. The mass range of the full-scan mass spectrum was 65-500 Da. The injection flow rates into the LCQ Advantage and Esquire-HPLC were 1 μL/min and 40 μL/min, respectively, through a post-column splitter. The instrument setup parameters for the LCQ Advantage were as follows: capillary voltage: 36 V; capillary temperature:150 °C; Tube lens offset 84 V; spray voltage:1.5kV; multipole 1 offset:-3.75 V; lens voltage:-16 V; multipole 2 offset:-7V; multiple RF amplitude:450 Vp-p. The instrument setup parameters for Esquire-HPLC were as follows: capillary voltage: 3500 V; endplate offset -500 V; nebulizer: 25 psi; dry gas: 9L/min; dry temperature:320 °C; skim 1:12 V; skim 2: 5V; cap exit offset: 10 V; octopole A: 2V; trap drive 65; oct RF 600 Vp-p; lens 1:-5 V; lens 2:-60 V; ICC target 10000. All experiments were run at 22 °C.

2.3. Results and Discussion

2.3.1. Principal of Operation on the Bifunctional HPLC Column

The Primsep 100 bifunctional column has a unique stationary phase, combining the properties of hydrophobic and ion-exchange stationary phases. Figure 2.1 shows the stationary phase consisting of negatively charged functional groups due to its embedded strong anionic long chain ion-pairing reagent. Strong interaction between the hydrophobic phase and ion-exchange stationary phase allows the control of the retention of ionic and neutral compounds independently. The developed applications of this stationary phase so far were focused on its mixed-mode condition, i.e., with an ion exchange gradient and an organic solvent composition gradient running simultaneously so that both the ionic and neutral compounds can be separated at once.[16-22] However, this column can also be run in single-mode separation. The column behaves like a reversed-phase HPLC column when a gradient of organic solvents is prepared under a specific buffering condition that can protonate the negatively charged functional groups. This can be done by applying high concentration of H^+ in the elution buffer so that the negatively charged functional groups on the stationary phase which are responsible for cation exchange separation mechanism are protonated by H^+. The workable pH range of the column is from pH 1.0 to pH 7. Most of the cation exchange separation function group was protonated when pH less than 1.0.

When a H^+ gradient is run from low to moderately high (pH is less than 1.5), the H^+ may not be sufficiently high enough to protonate all negative charges to switch off the cation exchange mode. In this case, the H^+ will compete with the positively charged analytes adsorbed on the column and elute them in the order of their ionic interaction with the negatively charged stationary phase from weak to strong, while the non-polar molecules will be retained on the stationary phase. The higher the net positive charge of an analyte, the stronger its ionic interaction with the negatively charged stationary phase is. As a result, they will be eluted from the column based on the cation exchange mechanism or the bifunctional column behaviors like a cation exchange HPLC column. If the solvent conditions used are in between the two extreme cases, a combination of the reversed phase and the cation exchange separations will be attained, resulting in a mixed-mode HPLC operation.

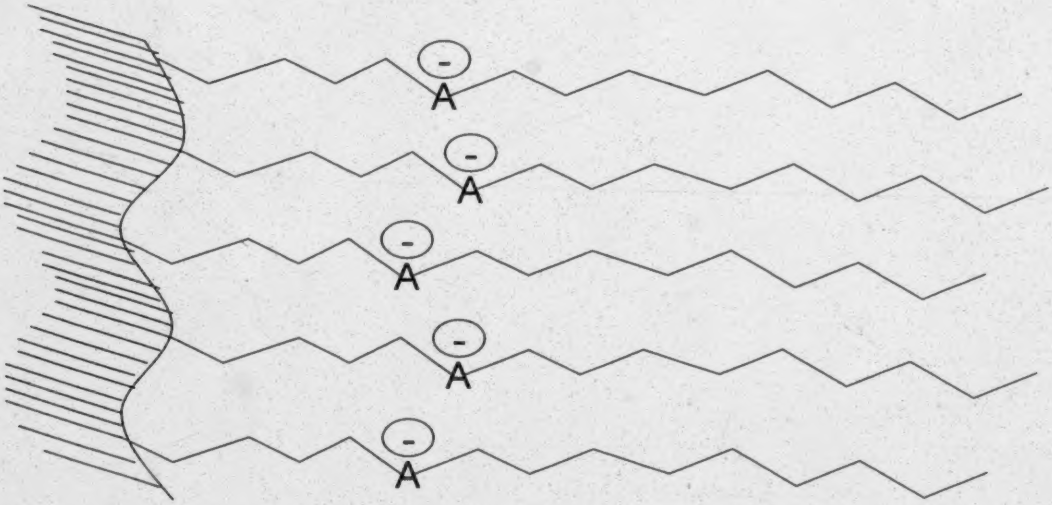


Figure 2.1. Schematic representation of the Primsep 100 bifunctional stationary phase. A^- represents the negatively charged functional group for cation exchange separation mechanism. The long alkyl chain on the right side of A^- is the group for reversed-phase separation mechanism. Details of the chemical structures are not available from the manufacturer.

2.3.2. Effects of Different Acids in the Mobile Phase on Separation

As discussed in the last section, concentration of H^+ is critical to control the separation mode of the bifunctional HPLC column. We can use acetic acid and TFA to illustrate the effects of different acids as H^+ provider in the mobile phase on separation. In this work, a 15 cm × 2.1 mm i.d. Primsep 100 column was used. A 30 min A-B linear gradient from 0% B to 100% B was applied. Buffer A was 1.5% acetic acid in 80:20 mixture of water-acetonitrile. The data are shown in Figure 2.2. The only difference in HPLC conditions between Figure 2.2 (A) and Figure 2.2 (B) was buffer B: 5% acetic acid with 0.1% TFA in 70:30 mixture of water-acetonitrile was used in Figure 2.2 (A) while 0.3% TFA in 70:30 mixture of water-acetonitrile was used in Figure 2.2 (B). 0.1% TFA added into buffer B in Figure 2.2 (A) was used to adjust the pH so that the pH of buffer B in (A) and (B) was the same (pH=1.4). Analytes 1 and 2 shown in Figure 2.2 were amino acid standards H and F, respectively. In this study H and F were selected because of their strong retention on the column due to their relatively high pI and hydrophobicity.

As Figure 2.2 shows, compared with (B), elution of H in (A) is delayed by more than 8 min and F can not even be eluted at all in (A). The pH of the two elution buffers is the same, but the elution strength of buffer B in (B) is much larger. This might be due to the fact that TFA is a much stronger ion-pair reagent than acetic acid in the mobile phase [29]. . The dissociated H^+ from TFA or acetic acid tends to bind the negatively charged stationary phase to replace the positively charged analytes, which are then released from the stationary phase. The H^+ acts as an eluent to take off the positively charged analytes. The complementary dissociated part of TFA or acetic acid, CF_3COO^- or CH_3COO^-, can associate with a positively charged analyte released from the stationary phase to form an ion-pair. On the other side the complementary CF_3COO^- or CH_3COO^- can act as a ion-pairing agent to hold the positively charged analytes, thus decreasing the possibility of the stationary phase recapturing the analytes. This "eluting-and-ion-pairing" process would assist in eluting the analytes from the column. The strength of the attacking agent H^+ is the same in Figure 2.2 (A) and (B). However, the ion-pairing ability of CF_3COO^- is much stronger than that of CH_3COO^-. As a consequence, the elution ability of TFA is much stronger than that of acetic acid.

This ion-pair theory can help the understanding of why the elution of H in Figure 2.2 (A) was much delayed and F could not even be eluted, though the pH of the elution buffer in (A) and (B) is the same. The fact that F could not be eluted was confirmed through its $M+H^+$ base peak. From Figure 2.2 it can be seen that the intensity of H in (B) is much stronger than in (A), even though buffer (B) has a much higher TFA concentration that can cause greater MS signal suppression [30]. A possible explanation for this inconsistency can also be based on the ion-pair theory. TFA, as a stronger ion-pair reagent, can form stronger ion-pairs with the analytes released from the stationary phase, resulting in an increase in the analytes' recovery from the column, which can then compensate TFA's signal suppression effect. The net result is a higher MS signal. Thus, the bifunctional column separation is affected by not only the concentration of H^+ but also the ion-pair ability of the running buffer. Because TFA's performance was found to be better than acetic acid, for the subsequent work, TFA, instead of acetic acid, was used as the acid for buffers A and B to provide H^+ for bifunctional column separation.

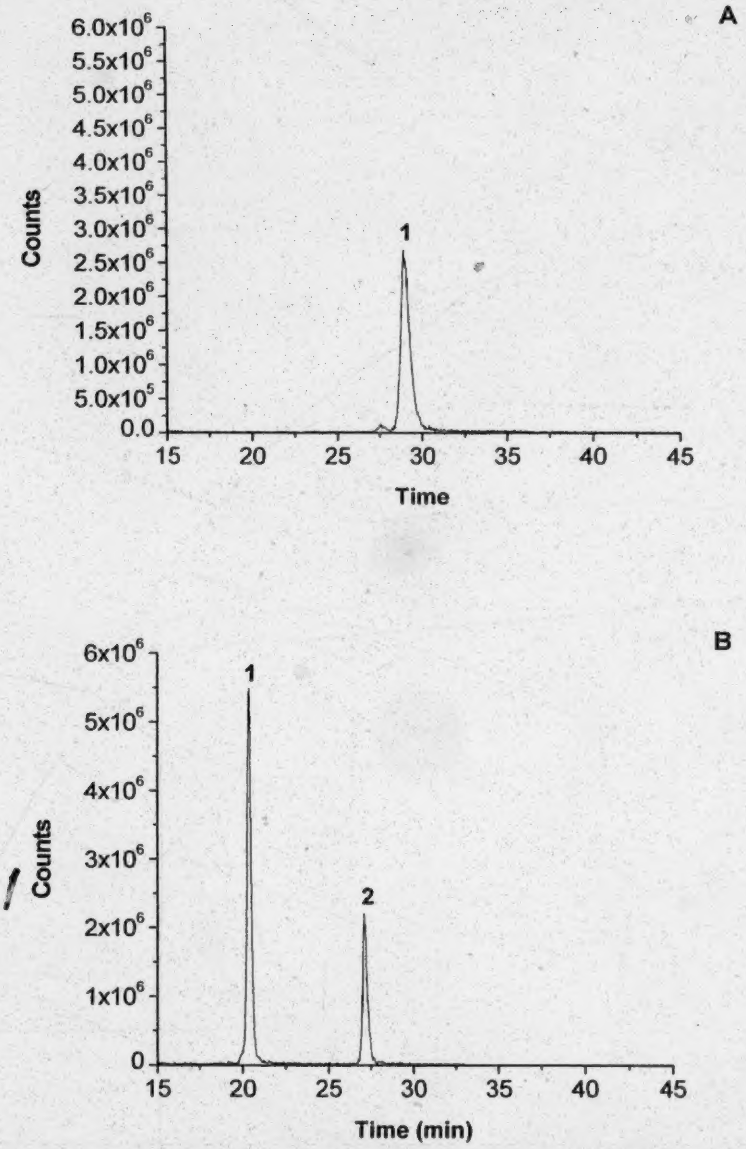


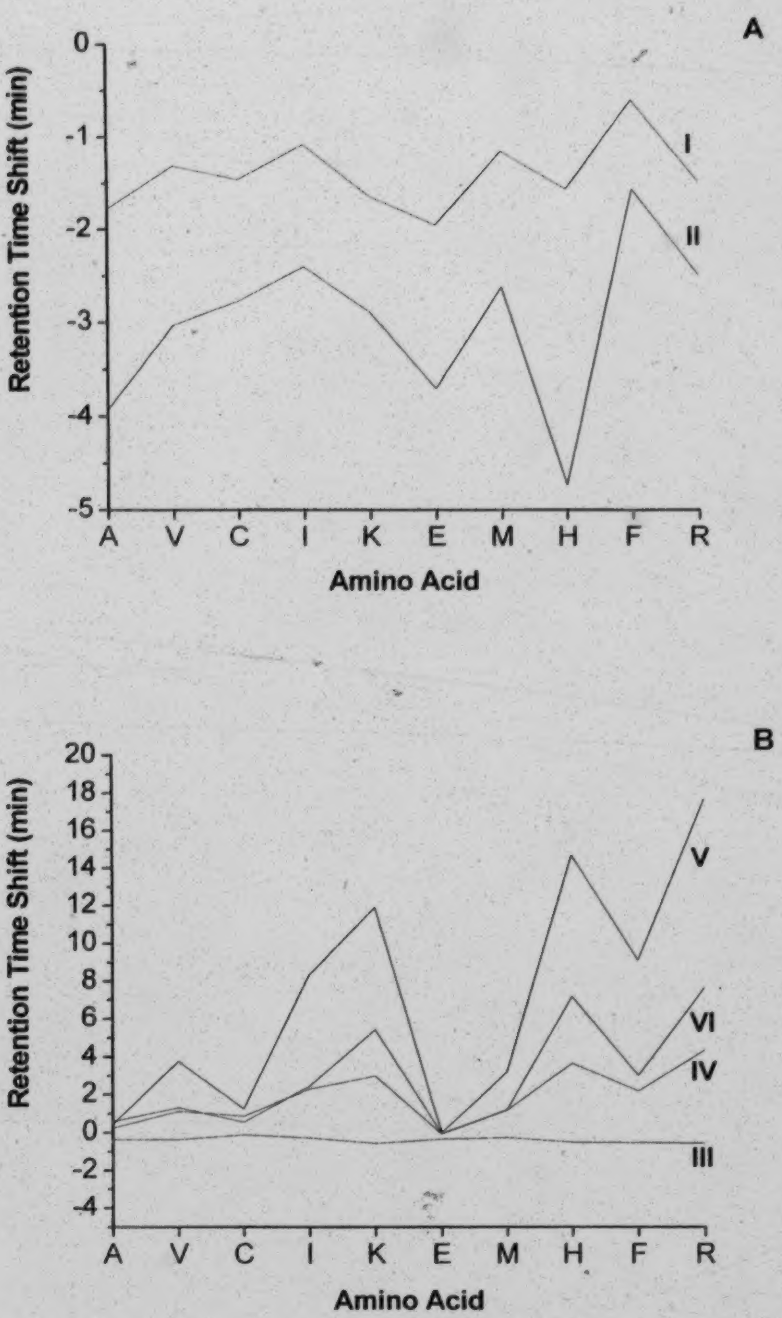
Figure 2.2. The M+H$^+$ base peaks of the mixed-mode HPLC/MS when(A) 5% acetic acid with 0.1% TFA in a 70:30 mixture of water-acetonitrile and (B) 0.3% TFA in a 70:30 mixture of water-acetonitrile was used as buffer B. A Primsep 15cm × 2.1mm i.d. column was run in a mixed-mode HPLC condition with a 30 min linear gradient. Other HPLC conditions for (A) and (B) were the same. The injection amount for each amino acid (1) H and (2) F was 2.5 nmol and 1.0 nmol respectively. The fact that F could not be eluted in (A) was verified through its M+H$^+$ base peak.

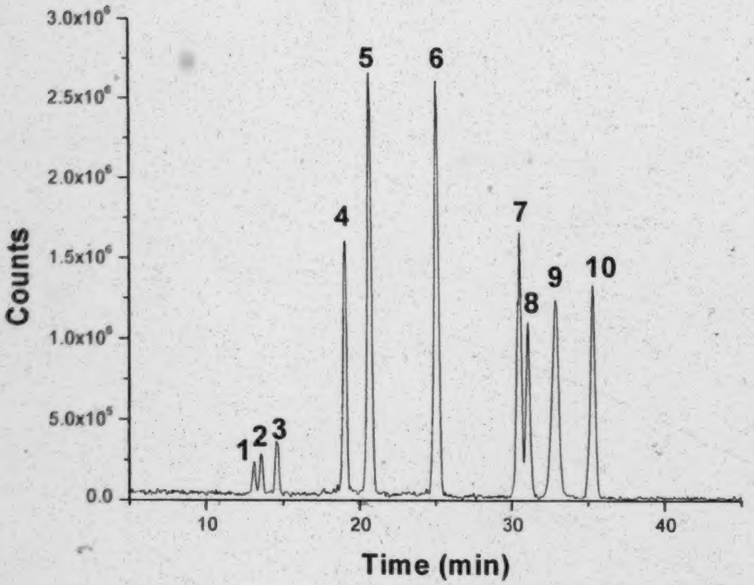
2.3.3 Effects of Long Chain Ion-Pair Reagents on Retention

Before ion-pair reagents were used to increase the retention of polar molecules and improve their separation on reversed-phase column [31], a similar idea had been practiced for many years: the long chain metal-chelate additives were extensively used as additives in the mobile phase to enhance the selectivity and separation efficiency of reversed phase HPLC [32, 33] and ion exchange HPLC [34] since 1978. In this study the effects of the long chain ion-pair reagents, HFBA, TFHA and HFNA on retention of amino acids on the bifunctional HPLC column have been studied. A 15 cm×2.1 mm i.d. Primsep 100 was run in the mixed-mode HPLC condition. A 45 min A-B linear gradient from 0% B to 100% B followed by 15 min 100% B was applied.

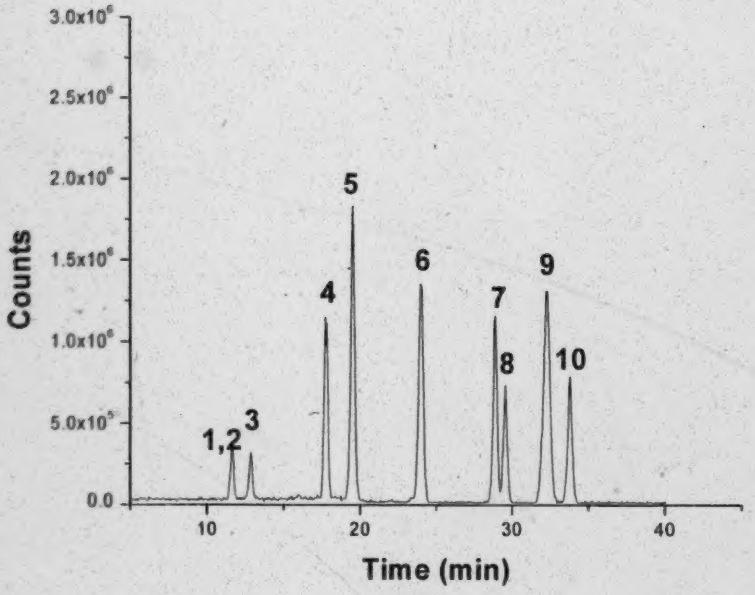
Figure 2.3 shows the retention time shifts of 10 amino acids: A, V, C, I, K, E, M, H, F, and R, when 5 mM HFBA (curve (I) in Figure 2.3 (A)), 10 mM HFBA (curve (II) in Figure 2.3 (A)), 1 mM TFHA (curve (III) in Figure 2.3 (B)), 5 mM TFHA (curve (IV) in Figure 2.3 (B)), 10 mM TFHA (curve (V) in Figure 2.3 (B)), or 1 mM HFNA (curve (VI) in Figure 2.3 (B)), respectively, was added to buffer A: 0.05% TFA in a 80:20 mixture of water-acetonitrile, and buffer B: 0.6%TFA in a 70:30 mixture of water-acetonitrile. In this study 10 amino acids were selected because their pIs were spaced in a wide range. Because the pKa of TFA is much less than the pKa of the long chain ion-pair reagents and the concentration of TFA is much greater than the concentration of the long-pair reagents, the influence of long-pair reagents on the pH of mobile phase could be ignored and the pH of the mobile phase was considered as constant when different long chain ion-pair reagents were added. The retention time shifts are calculated in relation to those determined under the same LC/MS conditions but without the use of any long chain ion-pair reagents. Curves (I)-(VI) in Figure 2.3 show that different long chain ion-pair reagents at different concentrations have different effects on the retention of the 10 amino acids. The curve (V) in Figure 2.3 (B) shows that HFNA ($C_9HF_{17}O_2$), the longest one of the three ion-pair reagents, tends to increase the analytes' retention even at a concentration in the mobile phase as low as 1 mM. However, TFHA ($C_7HF_{13}O_2$), which is shorter than HFNA, tends to increase the analytes' retention only when the concentration is greater than a critical concentration. Below its critical concentration, the ion-pair reagent actually reduces the analytes' retention ability. Above its critical concentration, an increase in TFHA concentration results in the larger reduction of the analytes' retentions. HFBA ($C_4HF_7O_2$), the shortest one of the three ion-pair reagents, tends to reduce the analytes' retention when its concentration is below 10 mM. The higher the concentration of HFBA, the more the analytes' retention times are reduced.

A simple model is put up as follows to explain these phenomena. In the mixed-mode HPLC condition, the functional group on the stationary phase is negatively charged and it can electrically repulse the negatively charged ion-pair ions, $HFNA^-$, $TFHA^-$, and $HFBA^-$, which are dissociated from the three ion-pair reagents in the mobile phase. On the other hand, the three negatively charged ion-pair ions can also bind the stationary phase through the intermolecular force, such as the dispersion force between the ion-pair reagents' hydrophobic chains and the hydrophobic group on the stationary phase. The longer the chain of the ion-pair reagent, the bigger the intermolecular force is. When the intermolecular force is big enough to overcome the repulsive ionic force between the negatively charged stationary phase and the ion-pair reagent, the ion-pair reagent can bind with the stationary phase. As a result, both the effective negative charge intensity and hydrophobicity are enhanced on the stationary phase, and then the retention of the positively charged analytes and non-polar analytes can be stronger. Conversely, if the intermolecular force cannot overcome the repulsive ionic force between the negatively charged station phase and the ion-pair reagent, the ion-pair reagent has to stay in the mobile phase. The negatively charged ion-pair reagent in the mobile phase can then compete with the negatively charged functional group on the stationary phase to pair the positively charged analytes. As a result, the retention of the analytes can become weaker. The chain of TFA is even shorter than HFBA, so its effect on retention of the analytes can be categorized into the same situation as HFBA.


A
Retention Time Shift (min)
0
-1
-2
-3
-4
-5
I
II
A V C I K E M H F R
Amino Acid
B
Retention Time Shift (min)
20
18
16
14
12
10
8
6
4
2
0
-2
-4
V
VI
IV
III
A V C I K E M H F R
Amino Acid


3.0x10⁶
2.5x10⁶
2.0x10⁶
1.5x10⁶
1.0x10⁶
5.0x10⁵
0.0
Counts
1
2
3
4
5
6
7
8
9
10
10
20
30
40
Time (min)

X


3.0x10⁶
2.5x10⁶
2.0x10⁶
1.5x10⁶
1.0x10⁶
5.0x10⁵
0.0
Counts
1,2
3
4
5
6
7
8
9
10
10
20
30
40
Time (min)

Y

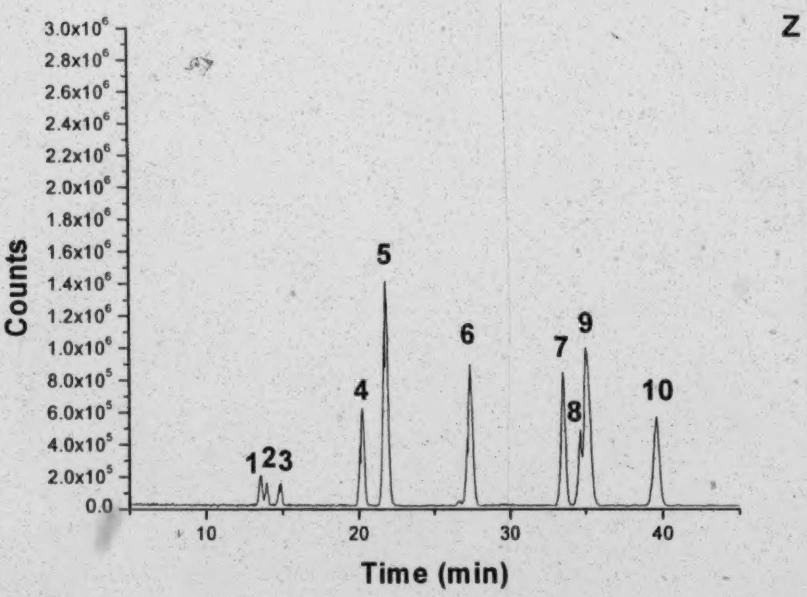


Figure 2.3. Retention time shift when (I) 5 mM HFBA; (II) 10 mM HFBA; (III) 1 mM TFHA; (IV) 5 mM TFHA; (V) 10 mM TFHA; and (VI) 1 mM HFNA was added to buffer A: 0.05% TFA in a 90:10 mixture of water-acetonitrile, and buffer B: 0.6%TFA in a 80:20 mixture of water-acetonitrile. The retention time shifts are calculated in relation to those determined under the same LC/MS conditions but without the use of any long chain ion-pair reagents. A Primsep 25 cm × 2.1 mm i.d. column was run in a mixed-mode condition with a 45 min linear gradient followed by 15 min 100% B. (X), (Y) and (Z) are base peaks chromatogram when no any long chain ion-pair reagents, 5 mM HFBA, and 5 mM TFHA was added respectively. 1 to 10 in (X), (Y) and (Z) stands for amino acids: C, E, A, V, M, I, K, H, F and R. The injection amount for each amino acid was 0.5 nmol.

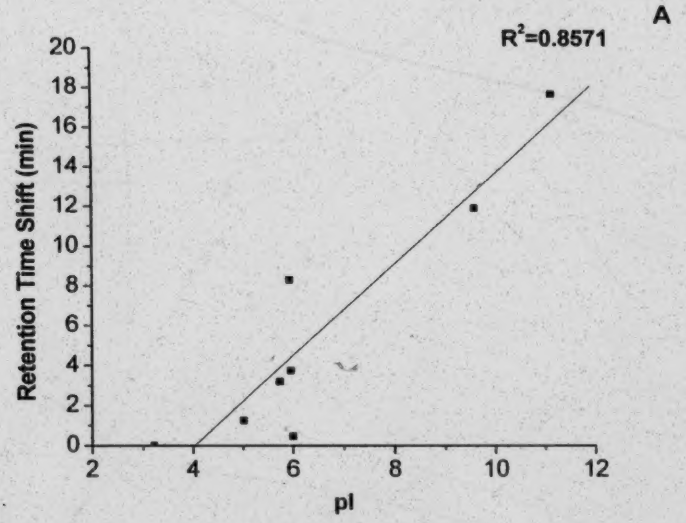
This explanation is consistent with the previous discussion regarding TFA's better performance than acetic acid as H^+ provider in the buffer in Section 2.3.2. To quantitatively analyzing the simple model we put forward, we use the equation from the linear solvent strength model:

$$\Delta t_a = \beta_a \times \Delta Q_s \tag{2.1}$$

, Δt_a is the retention time shift and ΔQ_s is the increase of ion exchange functional group density caused by ion-pair reagents adsorbing on the column, and β_a is the analyte's binding capacity with the ion exchange functional group [35]. When the concentration of an ion-pair reagent $C_{i,m}$ is below its critical concentration, the negatively charged ion-pair reagent in the mobile phase would compete with the negatively charged functional group on the stationary phase to pair the positively charged analytes, i.e., would shield the analyte's binding ability with the stationary phase, so the β_a in equation 2.1 is negative and decreases when $C_{i,m}$ increases. The higher the $C_{i,m}$ is, the more the β_a and Δt_a in equation 2.1 decreases, as shown in (I) and (II) in Figure 2.3. Conversely, when the concentration of an ion-pair reagent $C_{i,m}$ is above its critical concentration, the negatively charged ion-pair reagent tends to adsorb on the column and the negative charge density increases on the column, i.e., the ΔQ_s in equation 2.1 increases. In this case, the higher the $C_{i,m}$ is, the more the ΔQ_s and Δt_a in equation 2.1 increases, as shown in (IV) and (V) in Figure 2.3.

The extent of analyte charging is related to the pI of the molecule. The higher the pI, the higher the degree of positive ionization. Thus an analyte with a higher pI will bind stronger with the negatively charged functional group on the stationary phase, resulting in a bigger β_a in equation 2.1. Increasing β_a results in the increase of Δt_a in equation 2.1 when ΔQ_s is fixed. Thus the analytes' pI values should be proportionally related to their retention time shifts, if only ion exchange mechanism is in operation. Figure 2.4 shows the correlation between amino acid pI and their retention time shift. Because H and F are much more hydrophobic than the other 8 amino acids, due to their aromatic side chain, they are expected to display a large extent of reserved phase interaction. Thus they are not included in the pI-retention time shift correlation in Figure 2.4. Even though H and F are excluded for Figure 2.4, the correlation results shown in Figure 2.4 are still not very good. This suggests that even for the remaining 8 amino acids there are some degrees of reserved phase interaction present in the separation.

From Figure 2.3, it can be seen that the long chain ion-pair reagents' effect on the retention of the amino acid E is very limited. A possible reason is that amino acid E's pI is very low, i.e., close to 3. Its degree of ionization may not be as high as the other 9 amino acids. Thus, its β_a in equation 2.1 is much smaller. As a consequence, the effect of the long chain ion-pair reagents on its retention time shift is not as strong as on the other 9 amino acids. In the mixed-mode HPLC condition, it should be emphasized here that the reversed phase mechanism works with the cation exchange mechanism at the same time. Thus, the β_a in equation 2.1 is proportionally related to not only the analytes' pIs but also their hydrophobicity, and the ΔQ_s in equation 2.1 is proportionally related to not only the negative charge density increase on the stationary phase but also the hydrophobicity increase on the stationary phase.



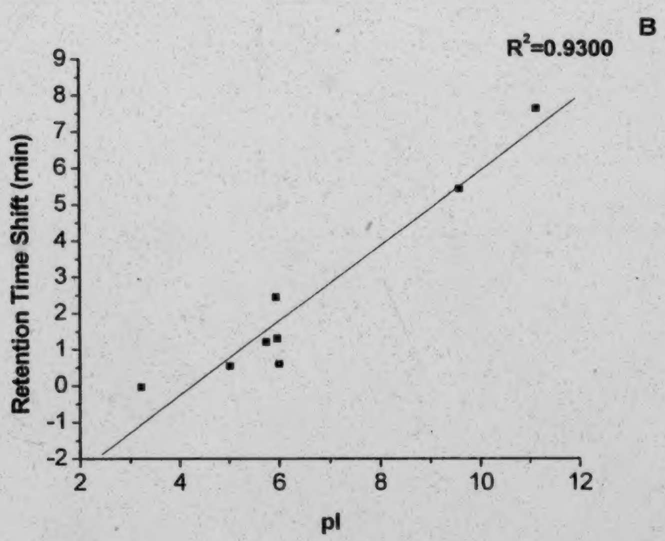


Figure 2.4. The correlation between pI of amino acids: A, V, C, I, K, Q, M, and R and their retention time shift with (A) 10 mM TFHA and (B) 1 mM HFNA. The raw data are the same as those for Figure 2.3.

Figure 2.5 shows the correlation between the retention time shifts when different long chain ion-pair reagents and concentrations were applied. The good correlations in

Figure 2.5 suggest that for this column the retention time shifts (Δt_a) in different buffer conditions are governed by the same mechanism related to the concentration and the type of the long chain ion-pair reagents. The longer ion-pair reagent with higher concentration tends to give greater retention times, when the concentrations of the long ion-pair reagents are above their critical concentration. As shown in Figure 2.5, (A) gives the best correlation, because this set of correlation data are from the same long chain ion-pair reagent but with different concentrations, while the correlation data in (B) and (C) are from different long chain ion-pair reagents with different concentrations. It is clear that the type of ion-pair reagent affects the correlation results.

An ion-pair reagent as long as TFHA, when it is used at a concentration greater than the critical concentration, can increase the retention. This interesting feature has been used to help separate amino acids.[31] However, our work described above illustrates that the retention enhancement varies widely for different analytes. This means that, for separating a complicated mixture, improvement in some compounds' separation can be obtained, but other compounds' separation may be worsened, due to inconsistency in retention enhancement. A likely situation suitable for applying ion-pair reagents is to separate several target compounds from a complex mixture sample. In this case, only the separation of a few specific compounds needs to be considered and separation of the other compounds can be ignored.

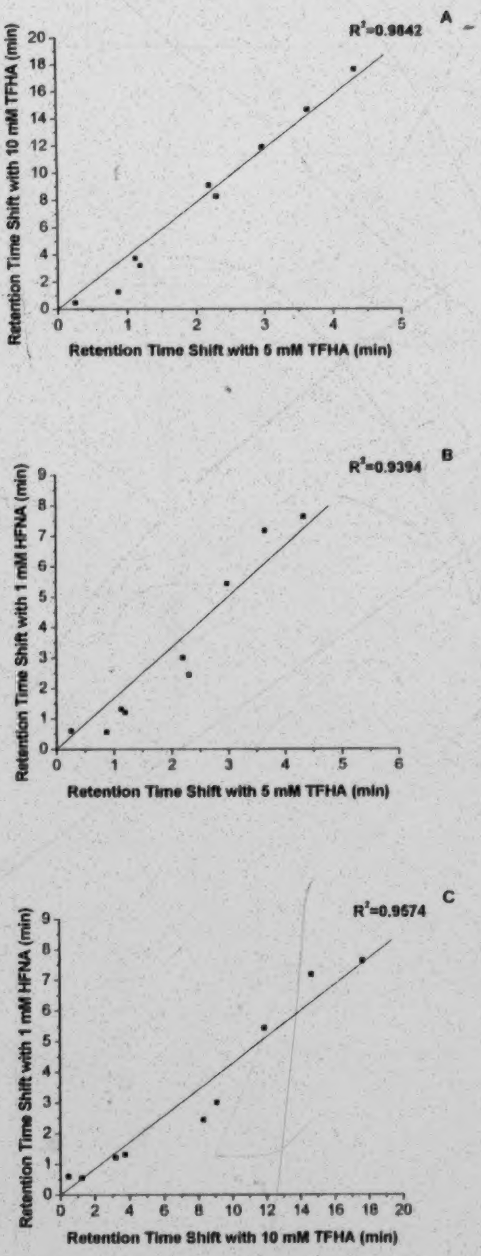


Figure 2.5. The correlation between retention time shifts when (A) 5 mM TFHA or 10 mM TFHA, respectively, applied; (B) 5 mM TFHA or 1 mM HFNA, respectively, was applied; (C) 10 mM TFHA or 1 mM HFNA, respectively, was applied. The raw data are the same as those for Figure 2.3.

Another disadvantage of using long chain ion-pair reagents is that they suppress analyte ion signals seriously in HPLC-MS.[36] The data shown in Table 2.1 illustrates the extent of ESI MS signal reduction when three long chain ion-pair reagents were used. In ESI MS, the negatively charged long chain ion-pair reagents preferentially adsorb at the liquid-air interface of the aerosol droplets formed in ESI. The cationic analytes are drowned and concentrated inside the aerosol droplets, so the surface concentration of analytes is decreased; thereby depressing the analyte MS signal. The longer the ion-pair reagent is, the more active it is on the surface of the aerosol droplets, and then, the more the MS signal is suppressed. The average of MS signal reduction with 5 mM HFBA, 5 mM TFHA, and 1 mM HFNA is found to be -31.9% , -47.6%, and -90%, respectively.

The third disadvantage of using long chain ion-pair reagents is that they can bind with the negatively charged functional group on the stationary phase through the intermolecular interaction between the long-chain ion-pair reagents and the stationary phase. As a result, much more time is required to regenerate the column after a run.[31]

After careful consideration of the pros and cons of using long chain ion-pair reagents for mixed mode separation using the bifunctional column, we abandoned this idea for separating amino acids and, instead, focused on a new method of performing different modes of HPLC in serial using the bifunctional column which is described below.

Table 2.1. Suppression of long chain ion-pair reagents on HPLC-ESI/MS signal of 10 amino acids. (-) means MS signal suppression. The raw data are the same as those for Figure 2.3.

	5 mM HFBA	5 mM TFHA	1 mM HFNA
C	-32.2%	-15.1%	-87.0%
E	-45.0%	-49.2%	-95.1%
A	-24.3%	-61.7%	-87.3%
V	-29.0%	-61.5%	-91.0%
M	-32.8%	-47.8%	-91.5%
I	-49.8%	-66.4%	-95.4%
K	-31.8%	-70.9%	-87.7%
H	-35.4%	-25.5%	-73.2%
F	2.7%	-20.7%	-94.1%
R	-41.7%	-57.4%	-95.8%

2.3.4. Running the Bifunctional Column in Single Separation Mode: Reversed Phase or Ion Exchange

From the very beginning, the bifunctional column was proposed to be used for separating the analytes that are too close in physical or chemical characteristics to be separated by conventional columns.[15] All the applications developed so far were designed to run the columns in mixed-mode HPLC.[16-22] These successful applications demonstrate that a mixed-mode HPLC is a powerful technique. In our work, however, we develop a method of running the bifunctional column in reversed phase, mixed-mode, and ion exchange separation in series, instead of running it in mixed mode only. In this section, two simple mixtures are separated to illustrate the advantages of running single mode separation over mixed mode separation.

In this case, a 15 cm × 2.1 mm i.d. Primsep 100 was used. A 30 min A-B linear gradient from 0% B to 100% B followed by 15 min at 100% B was applied. Three amino acids, namely F, H, and K, were used to evaluate the separation performance of the column performed in mixed-mode and reversed phase HPLC conditions, respectively. The buffers for the mixed-mode HPLC conditions were: 0.05% TFA (buffer A) and 0.5% TFA (buffer B), respectively, in a 75:25 mixture of water-acetonitrile. The buffers for the reversed-phase HPLC conditions were: 0.3% TFA in 95:5 mixture of water-acetonitrile (buffer A) and 0.3% TFA in a 20:80 mixture of water-acetonitrile (buffer B). In reversed phase HPLC condition, the column was first equilibrated with buffer A that contained a high concentration of H^+ and a very low concentration of organic solvents. The high concentration of H^+ was used to block the negatively charged cation exchange functional group on the column so that principally only the reversed phase mechanism was effective.

Figure 2.6 (A) shows the $M+H^+$ base peak chromatogram obtained under the mixed-mode HPLC condition, while (B) shows the $M+H^+$ base peak chromatogram obtained under the reversed phase-mode HPLC condition. 1, 2, and 3 correspond to the amino acids F, H, and K, respectively. Figure 2.6 (A) shows that the three analytes co-elute under the mixed-mode HPLC condition, while their separation in the reversed phase HPLC condition, shown in Figure 2.6 (B), is significantly improved. The results can not exclude the possibility of separating the three amino acids under other mixed-mode HPLC conditions by painstakingly adjusting the concentration of H^+ and the organic composition. But the results do demonstrate that using the reversed phase condition the three amino acids can be readily separated without much effort on method development.

The pI of the three amino acids are F < H < K. Based on the cation exchange separation mechanism, F would be the first and K would be the last to be eluted. The reason is that, as already discussed in section 2.3.3, the higher the pI is, the higher the degree of the positive ionization, and the stronger the analyte's binding with the negatively charged functional group on the stationary phase. However, the hydrophobicity of the three amino acids are K < H < F. Based on the reversed phase separation mechanism, F would be the last and K would be the first to be eluted. It can be seen that in this case the two separation mechanisms counteract each other. The overall effect of the cation exchange mechanism and the reversed phase mechanism was that the three amino acids could not be separated in the mixed-mode HPLC conditions described.

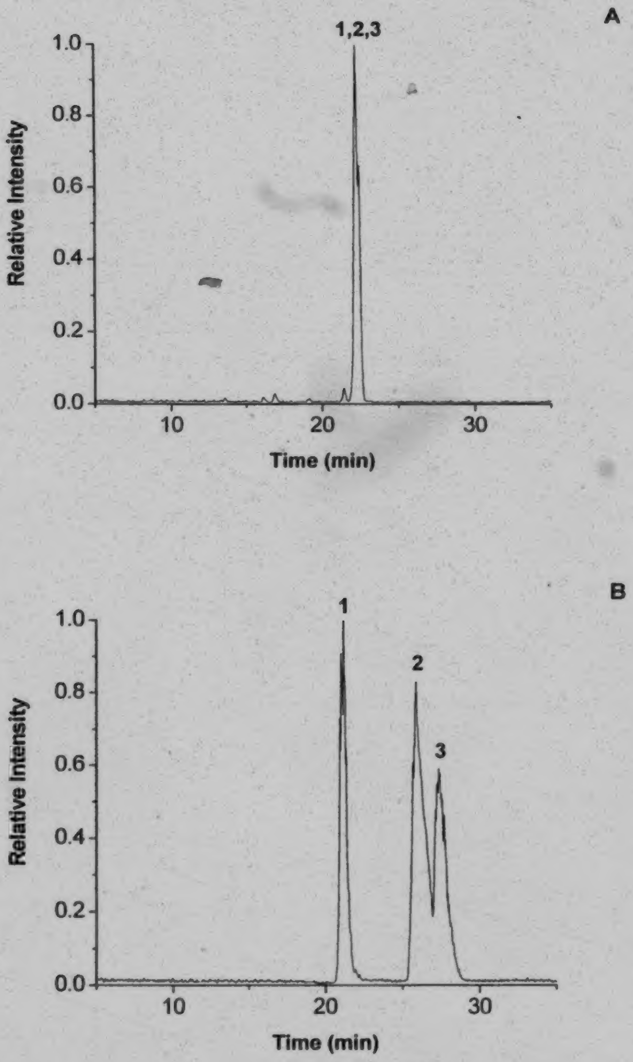


Figure 2.6. The M+H$^+$ base peaks of: (A) mixed-mode HPLC/MS and (B) reversed phase HPLC/MS used to separate amino acids: (1) F, (2) H, and (3) K. A Primsep 100 25cm × 2.1 mm i.d. column was used in both modes. A 30 min A-B linear gradient from 0% B to 100% B followed by 15 min at 100%B was applied. The buffers for the mixed-mode HPLC conditions were: 0.05% TFA (buffer A) and 0.5% (buffer B), respectively, in a 85:15 mixture of water-acetonitrile. The buffers for the reversed-phase HPLC conditions were: 0.3% TFA in 95:5 mixture of water-acetonitrile (buffer A) and 0.3% TFA in 20:80 mixture of water-acetonitrile (buffer B). Injections contained 2.5 nmol of each amino acid.

Next, we ran the column in cation exchange HPLC and compare it with the mixed-mode HPLC. A 30 min A-B linear gradient from 0% B to 100% B followed by 15 min at 100%B was applied in both modes. The buffers for the cation exchange HPLC conditions were: 0.05% TFA in 95:5 mixture of water-acetonitrile (buffer A) and 0.3% TFA (buffer B), respectively, in a 95:5 mixture of water-acetonitrile. The buffers for the mixed-mode HPLC conditions were: 0.05% TFA (buffer A) and 0.5% (buffer B), respectively, in 80:20 mixture of water-acetonitrile. The results are shown in Figure 2.7, where analytes 1 and 2 represent amino acids C and N, respectively. Figure 2.7 (A) shows the chromatogram obtained with the mixed-mode HPLC condition, while Figure 2.7 (B) shows the chromatogram obtained with the cation exchange mode HPLC condition. The resolution of C and N (R_s) in Figure 2.7 (A) is 0.3, while their resolution R_s in Figure 2.7 (B) is 1.4. The pI of C is smaller than N. Based on the earlier discussion, C would be the first and N would be the second to be eluted, when the cation exchange separation mechanism is in operation. On the other hand, the hydrophobicity of the 2 amino acids are N < C. Based on the reversed phase separation mechanism, C would be the second and N would be the first to be eluted. The overall effect of the two counteracting cation exchange mechanism and the reversed phase mechanism is that the separation of the 2 amino acids in mixed-mode HPLC is not as good as that in cation exchange HPLC only.

We can conclude from the two examples described above that the bifunctional HPLC column has versatility over a regular single-mode HPLC column since it can be run in different HPLC modes without necessarily switching the columns and this versatility provide us flexibility for method development. In addition, a mixed-mode separation mechanism is not necessarily better than a single-mode separation mechanism in terms of the separation resolution and the complexity of method development (e.g., both pH and solvent composition gradients need to be optimized). If the cation exchange mechanism and reversed phase mechanism in mixed-mode separation neutralize each other, the separation factors will actually be reduced.

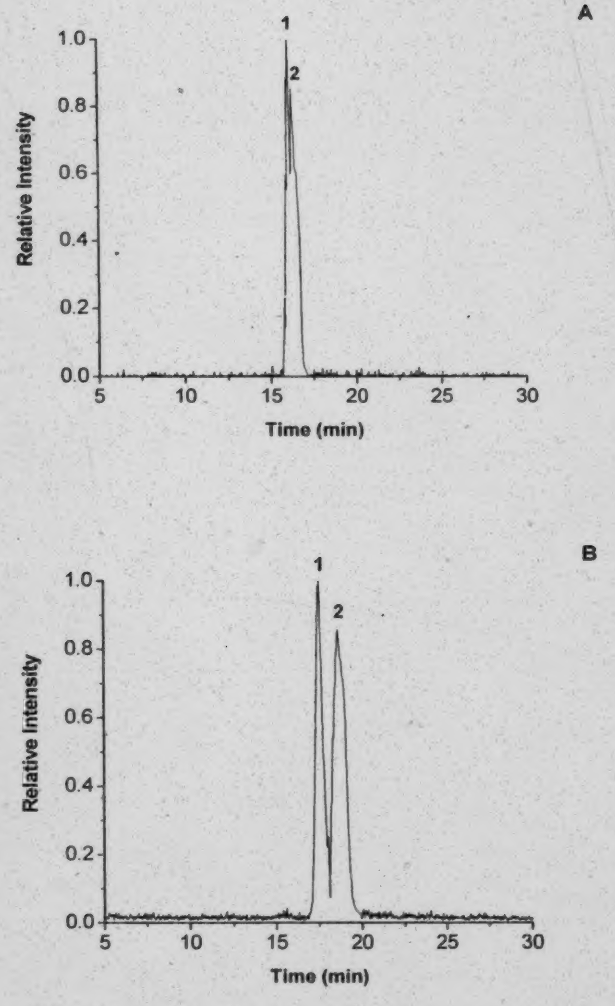


Figure 2.7. The M+H$^+$ base peaks of: (A) mixed-mode HPLC/MS and (B) cation exchange HPLC/MS used to separate amino acids (1) C and (2) N on a Primsep 100 25 cm × 2.1 mm i.d. mixed-mode HPLC column. A 30 min A-B linear gradient from 0% B to 100% B followed by 15 min at 100% B was applied in both modes. The buffers for the cation exchange HPLC conditions were: 0.05% TFA in 95:5 mixture of water-acetonitrile (buffer A) and 0.3% (buffer B) respectively in a 95:5 mixture of water-acetonitrile. The buffers for the mixed-mode HPLC conditions are: 0.05% TFA (buffer A) and 0.5% (buffer B) respectively in a 80:20 mixture of water-acetonitrile. Injections contained 2.5 nmol of each amino acid.

2.3.5. Multidimensional Separation Using a Bifunctional HPLC Column and Its Application in Urinanalysis

As illustrated in last section, the bifunctional HPLC column can be used to perform different modes of separation. In this section, we demonstrate that this column can be run in ion exchange mode, mixed-mode, and reversed phase mode in series to form a multidimensional separation on a single column with one HPLC system. The column is run in 3 modes rather than in one mixed-mode only.

In this experiment, a 25 cm × 2.1 mm i.d. Primsep 100 was used. A 25 min A-B linear gradient from 0% B to 100% B for the cation exchange separation, followed by a 15 min B-C linear gradient from 100% B to 100% C for mixed-mode separation, and then a 40 min C-D linear gradient from 100% C to 100% D for the reversed phase separation. Buffer A was 0.05% TFA in 92:8 mixture of water-acetonitrile. Buffer B was 0.1% TFA in a 92:8 mixture of water-acetonitrile. Buffer C was 0.3% TFA in a 80:20 mixture of water-acetonitrile. The buffer D was 0.3% TFA in a 20:80 mixture of water-acetonitrile.

Although better than long chain ion-pair reagents like HFNA, TFA still suppresses HPLC/MS signals, resulting in the reduction of overall detection sensitivity. Therefore, before combining the bifunctional column separation with ESI MS, we need to address this suppression issue. In 1995, a method, called TFA-fix, was developed to reduce TFA's suppression of ESI-MS signals and improve the sensitivity of peptide mapping [37]. The basic idea of TFA-fix was to apply post-column organic acids to change the ion-pair formation and organic solvent to modify the composition of the droplets formed in ESI. This approach was applied to overcome the TFA suppression in our multidimensional separation method. 2-propanol was used since it has been proved to be the best for TFA-fix.[37] Through a tee, a 25:75 mixture of organic acid:2-propanol mixture at a flow rate of 0.067 mL/min was mixed with the eluate from the HPLC column. The schematic of the analytic setup for this section is shown in Figure 2.8. Formic acid, acetic acid and propionic acid, respectively, were used to form organic acid/2-propanol mixtures to evaluate their performance as post-column TFA-fix additives.

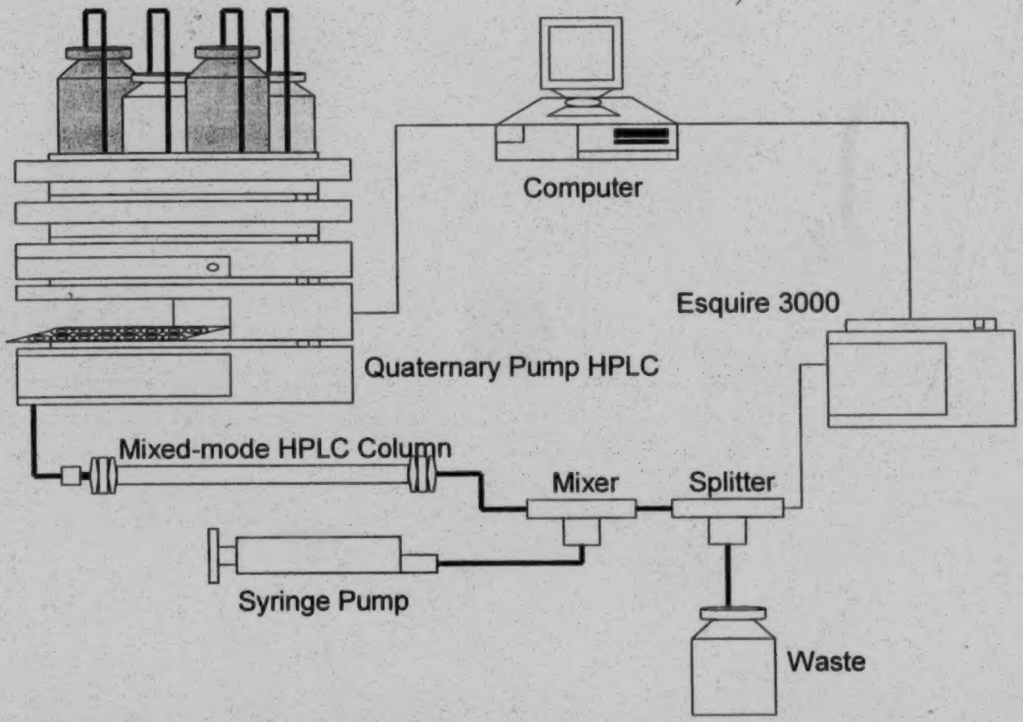


Figure 2.8. The schematic of the analytical setup for section 2.3.5. A 25 cm × 2.1 mm i.d. Primsep 100 was used in different modes. A 25 min A-B linear gradient from 0% B to 100% B for the cation exchange separation followed by a 15 min B-C linear gradient from 100% B to 100% C for mixed-mode separation and then a 40 min C-D linear gradient from 100% C to 100% D for the reversed phase separation. Buffer A was 0.05% TFA in 92:8 mixture of water-acetonitrile. Buffer B was 0.1% TFA in 92:8 mixture of water-acetonitrile. Buffer C was 0.3% TFA in 80:20 mixture of water-acetonitrile. Buffer D was 0.3% TFA in 20:80 mixture of water-acetonitrile.

Figure 2.9 shows the base peak chromatogram of the amino acids: C, Y, V, K, W and TR, where (A) is without; and (B) is with post-column acetic acid/2-propanol as the TFA-fixer. On average, the acetic acid/2-propanol TFA-fixer can enhance the MS signal by 740%. Figure 2.10 shows the performance of other post-column additives as the TFA-fixer compared with the acetic acid/2-propanol mixture. The data show that if only 2-propanol was used as the post-column additive, the MS signal could also be enhanced by 30% on average. A possible mechanism of 2-propanol as the TFA-fixer is that 2-propanol could reduce the surface tension [37], a tangential force that keeps a fluid together at the air/ESI droplet interface. Like the long chain ion-pair reagents, TFA might distribute on the surface more than inside the ESI droplets. Volatile TFA might benefit from the surface tension reduction and escape more easily into the air, and the TFA concentration in the surface of ESI droplets would be reduced, so the suppression from TFA would be reduced. The data in Figure 2.10 show that the MS signal enhancement by 2-propanol alone is much less than the enhancement with the acetic acid/2-propanol mixture. This means that the organic acid's ion-pair mechanism, shown in the equations 2.2 - 2.4, might play a more important role than the organic solvent's surface tension reduction:

$$MH^+ \bullet CF_3COO^- \overset{\rightharpoonup}{\longleftarrow} MH^+ + CF_3COO^- \qquad (2.2)$$

$$CF_3COO^- + RCOOH \overset{\longleftarrow}{\longrightarrow} RCOO^- + CF_3COOH \uparrow \qquad (2.3)$$

$$MH^+ + RCOO^- \overset{\longleftarrow}{\longrightarrow} MH^+ \bullet RCOO^- \qquad (2.4)$$

Summing up from the equation 2.1 to equation 2.3 gives equation 2.5:

$$MH^+ \bullet CF_3COO^- + RCOOH \overset{\longleftarrow}{\longrightarrow} MH^+ \bullet RCOO^- + CF_3COOH \uparrow \qquad (2.5)$$

, where RCOOH stands for the organic acid. Due to the strong ion-pair ability of CF_3COO^-, reaction 2.2 favors its reverse reaction. In terms of Brønsted theory, an acid-base reaction always favors the equilibrium towards the weaker acid and the weaker base. Compared to CF_3COOH and $RCOO^-$, RCOOH and CF_3COO^- are, respectively, weaker acids and weaker bases. It means that, based on the acid-base theory, reaction 2.3 also favors its reverse reaction. However, if CF_3COOH is more volatile than the organic acid RCOOH, CF_3COOH would be more easily brought into the gas phase and the resulting entropy increment could facilitate the reaction 2.3. The TFA-fix effect increased in the order: formic acid, acetic acid, and propionic acid, which also follows the order of boiling points of the 3 organic acids. The higher the organic acid's boiling point is, the bigger the entropy increment in reaction 2.3, and then the more likely reaction 2.3 can take place.

The other factor for MS signal enhancement is related to the disassociation ability of $MH^+ \cdot RCOO^-$. After the product CF_3COOH of the reaction 2.5 is evaporated into the gas phase, the other product $MH^+ \cdot RCOO^-$ may stay on the surface of the ESI droplets. $MH^+ \cdot RCOO^-$ is much more easily disassociated than $MH^+ \cdot CF_3COO^-$. The reason is that intra-molecular interactions, such as permanent dipole-induced dipole interactions and permanent dipole-permanent dipole interactions between the ionic MH^+ and $RCOO^-$, are much smaller than those between the ionic MH^+ and CF_3COO^-, since the polarity of $RCOO^-$ is much smaller than that of CF_3COO^-. The more likely $MH^+ \cdot RCOO^-$ is disassociated, then the more likely MH^+ can enter into the gas phase from the ESI droplets driven by the repulsive coulombic interaction between MH^+, therefore the higher the MS signal is. The non-polarity of the organic acids increases in the order: formic acid, acetic acid, and propionic acid, so the disassociation ability of $MH^+\cdot RCOO^-$ also increases in this order. Because both the entropy increment facilitating reaction 2.3 and the disassociation ability of $MH^+\cdot RCOO^-$ increases from formic acid to acetic acid to propionic acid, the MS signal enhancement should increase in the same order. This agrees with the obtained data shown in Figure 2.10. In the following study acetic acid/2-propanol, instead of the propionic acid/2-propanol mixture, was used as the post-column TFA-fixer to avoid the very strong odor smell from propionic acid.

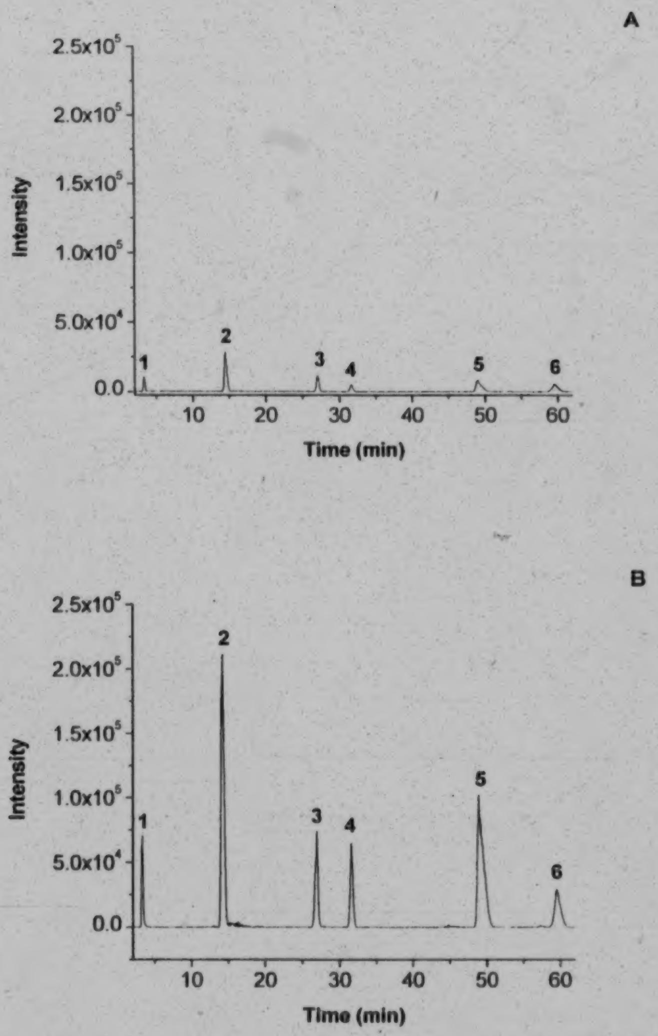


Figure 2.9. The base peak chromatogram of TR (1), amino acids C (2), Y (3), V (4), K (5) and W (6). (A) without and (B) with a 25:75 mixture of acetic acid/2-propanol at a flow rate of 0.067mL/min was used as the post-column TFA-fixer. Other HPLC conditions are shown in Figure 2.8. The injection for analytes 1-6 was 3.2, 8.5, 5.0, 3.4, 12.0, 0.2 nmol respectively.

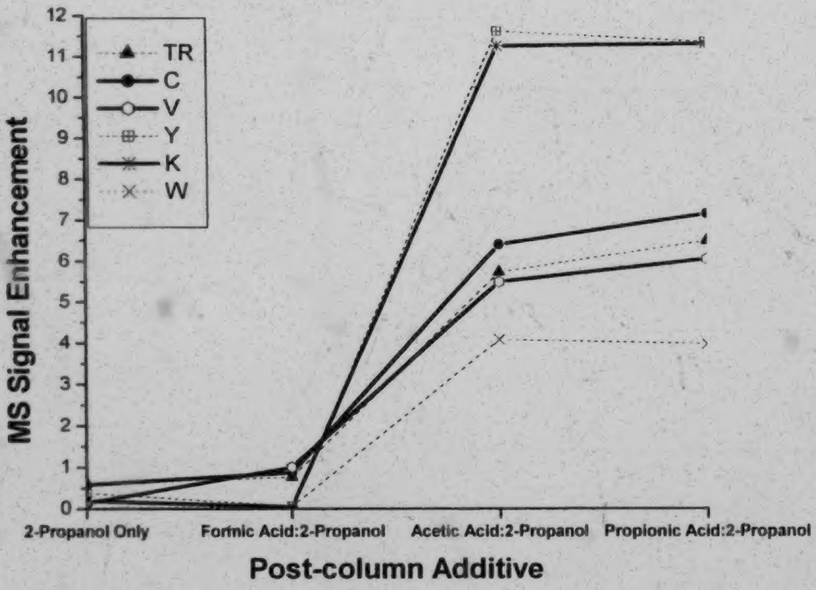


Figure 2.10. The performance of other post-column additives as the TFA-fixer contrasted with the acetic acid/2-propanol mixture. Except for the different post-column additives shown along the x-axis, all other HPLC conditions were the same as those in Figure 2.9.

Analysis of the amino acids in human body fluids is a crucial part of the clinical diagnosis of metabolism disorders, and in nutritional assessment.[38] Both reversed phase and ion exchange method are used in clinical laboratories to separate amino acids. Due to the high polarity of the amino acids, all of the developed methods with reversed phase columns had to use derivatization [39, 40], alkylsulfonates [33] or a long-chain ion-pair reagent [31] to increase their non-polarity so that the amino acids would have increased retention on the reversed phase columns. The ion exchange methods with [38] or without derivatization [41] were developed for clinical analysis of amino acids. Generally, ion exchange methods are more popular with better validations in clinical analyses. Ion-exchange chromatography with UV detection by means of the ninhydrin reaction is now the most commonly used method to determine amino acid concentrations in biological fluids and is considered as the reference method. Although its HPLC has been fully automated, this method generates results which are generally unsatisfactory in various quality controls, mainly related to unskilled users running the derivatization [38]. An ion exchange method without derivatization developed by Dionex (Sunnyvale, CA, USA), which demonstrated its capacity of separating 17 free amino acids, can appreciably overcome this deficiency and has been widely used in food and drug analysis.[41-43]

To demonstrate our method's application in complex mixture separation, 20 amino acids and 4 of their metabolites in humans, CS, TR, OT, and CT, were separated. Figure 2.11 is the base peak chromatogram of the 24 analytes. Peaks 1-24 are: TR, D, C, N, S, Q, T, P, E, G, A, V, M, CT, Y, I, L, F, OT, H, K, R, CS and W, respectively. The injection amount of TR, D, C, N, S, Q, T, P, E, G, A, V, M, CT, Y, I, L, F, OT, H, K, R, CS and W was 1.3, 0.6, 2.8, 1.7, 0.5, 0.6, 1.4, 0.4, 1.1, 4.5, 9.7, 1.3, 1.4, 2.1, 0.8, 0.9, 0.9, 0.4, 3.1, 1.6, 1.3, 3.4, 2.2, 0.05 nmol, respectively. Except the separation between Q and T in cation exchange mode, all the other separation factors were > 1.2. A possible reason for the poor separation between Q and T is that the pI difference between Q and T is as small as 0.01 and this slight difference causes the difficulty to separate them in ion exchange HPLC. To our knowledge, this is the best separation method of the whole 20 amino acids that has ever been reported with MS-compatible HPLC conditions.

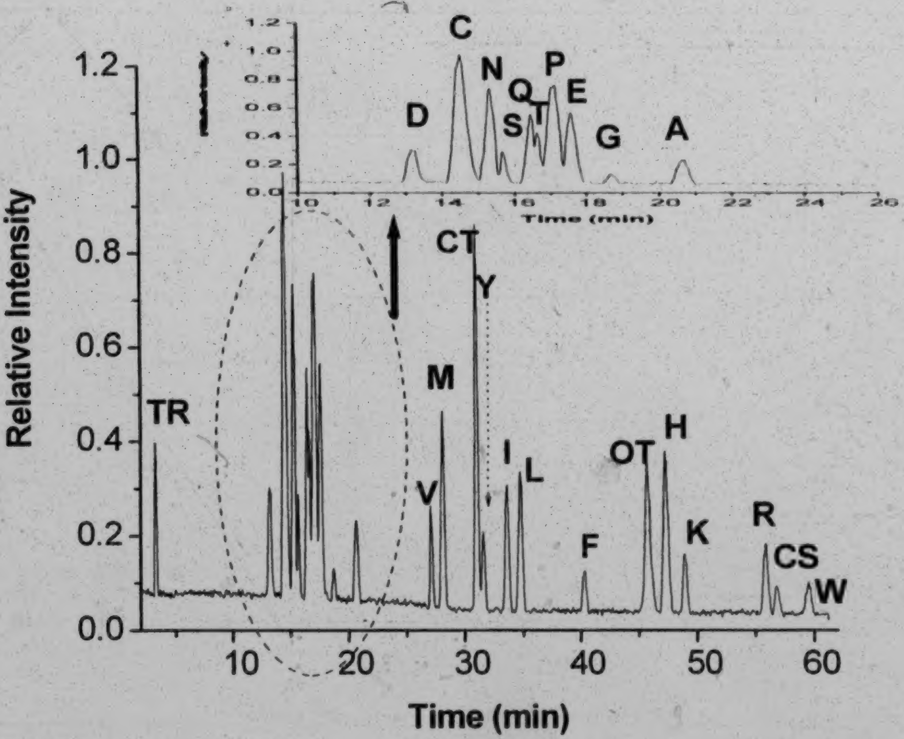


Figure 2.11. The base peak chromatogram of the 24 analytes separated by multidimensional HPLC separation on one single mixed-mode column. The injection amount of TR, D, C, N, S, Q, T, P, E, G, A, V, M, CT, Y, I, L, F, OT, H, K, R, CS and W are: 1.3, 0.6, 2.8, 1.7, 0.5, 0.6, 1.4, 0.4, 1.1, 4.5, 9.7, 1.3, 1.4, 2.1, 0.8, 0.9, 0.9, 0.4, 3.1, 1.6, 1.3, 3.4, 2.2, and 0.05 nmol, respectively. Other HPLC conditions are shown in Figure 2.8.

The analytes separated in cation exchange mode (TR, D, C, N, S, Q, T, P, E, G, A) are very small and polar compounds, and the analytes separated in reversed phase mode are those bigger ones showing some non-polarity. The good correlation between the MS signal intensity of $M+H^+$ (except for CT that itself is in ionic form, so M^+ was used instead) and the concentrations shown in Table 2.2 means that the developed multidimensional method can also be used for quantification. Figure 2.12 shows the correlation graph for (I) C and (II) D separated in cation exchange mode, and (III) CT and (IV) L separated in mixed-mode, and (V) CS and (VI) K separated in reversed phase mode.

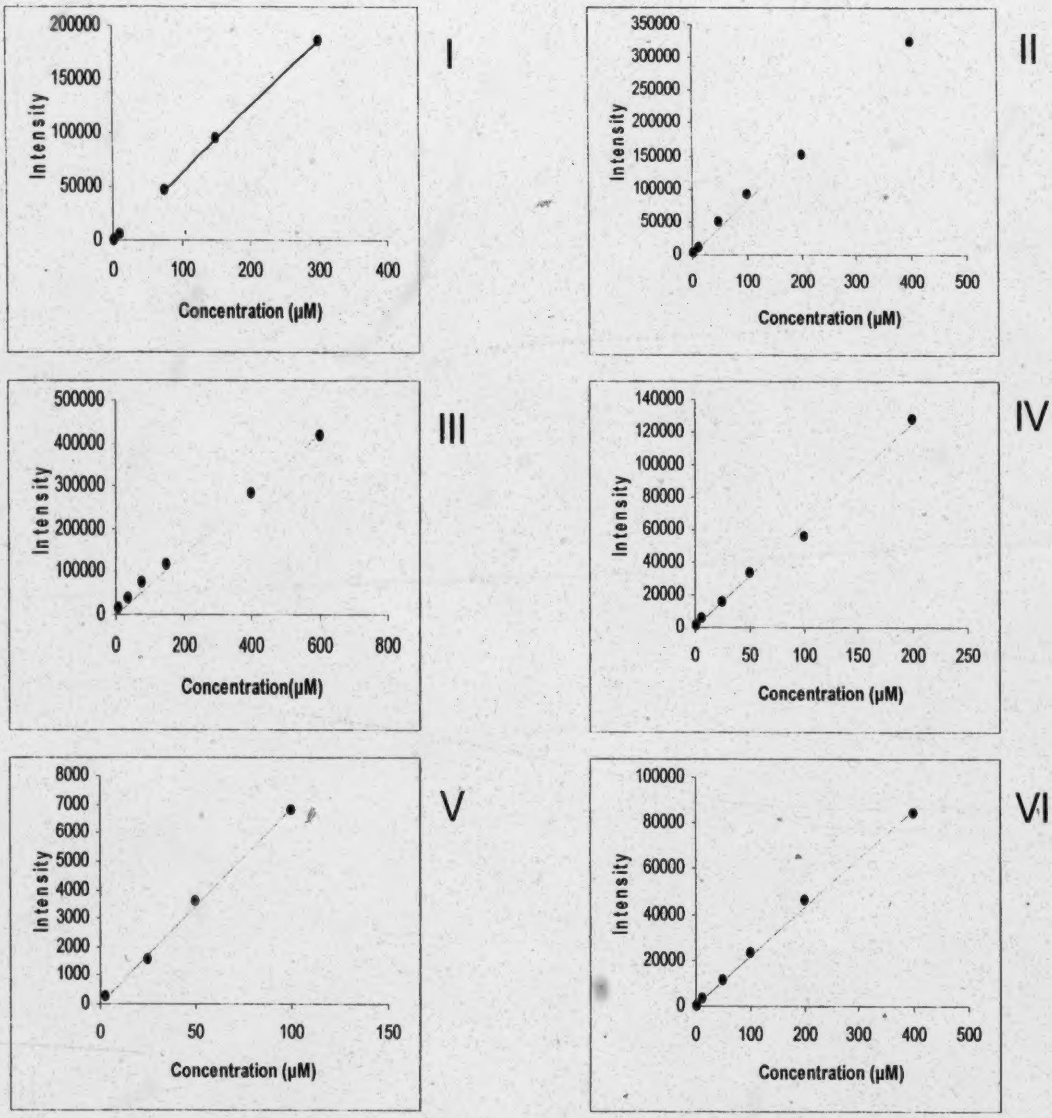


Figure 2.12 The correlation graph for (I) C and (II) D separated in cation exchange mode, and (III) CT and (IV) L separated in mixed-mode, and (V) CS and (VI) K separated in reversed phase mode. The raw data are the same as those for Table 2.2.

Table 2.2. Summary of R^2 and the linear range of MS signal intensity-concentration correlations, and LODs of the 24 analytes listed in Figure 2.11. HPLC conditions are shown in Figure 2.8.

Analytes	Retention Time (min)	R^2	Linear Range (<μM)	LOD (μM)	Average Concentration in the Sampled Urine (n=3)	CV (%) of the Calculated Concentration in the Sampled Urine (n=3)
TR	3.4	0.9931	150	0.46	174.4	2.13
D	13.1	0.9962	400	0.78	13.2	1.36
C	14.5	0.9999	300	0.29	39.0	1.42
N	15.2	0.9973	600	1.17	428.3	0.81
S	15.7	0.9933	100	1.55	278.0	1.01
Q	16.4	0.9815	100	0.48	617.5	5.05
T	16.7	0.9906	600	1.17	462.9	4.15
P	17.1	0.9935	100	0.26	27.8	3.73
E	17.6	0.9928	500	0.78	1109.6	2.82
G	18.7	0.9989	1000	7.30	1689.4	1.33
A	20.7	0.9942	1200	9.25	1476.6	2.14
V	27.1	0.9932	300	1.60	1208.2	1.65
M	28.1	0.9995	300	0.58	33.6	3.14
CT	31.0	0.9939	600	0.58	638.7	3.25
Y	31.7	0.9944	100	0.46	319.7	1.38
I	33.6	0.9922	200	0.39	20.8	2.59
L	34.8	0.9946	200	0.39	50.2	1.94
F	40.4	0.9969	100	0.39	110.2	2.73
OT	45.8	0.9921	600	1.17	36.5	1.67
H	47.3	0.9934	400	0.39	1624.7	1.65
K	48.9	0.9999	400	0.78	164.2	2.34
R	55.9	0.9920	300	0.58	111.0	1.06
CS	57.0	0.9980	100	0.78	14.1	1.9
W	59.5	0.9917	100	0.23	10.2	2.99

All of the R^2 in Table 2.2 are >0.991, except for Q and T. One possible reason is the relatively poor separation between Q and T, as shown in Figure 2.11. The data show how important separation is for quantitative analysis with all detectors like ion trap MS that has ion-ion repulsion and space-charge issue limiting its resolution, sensitivity, and trapping capacity. In fact, the signal interference, either weaker or stronger, caused by the coexisting analytes is present in all kinds of analytical instruments. Amino acid quantitation without any separation was carried out by NMR [44], but no wide range of applications can validate such kind of quantitative analysis method without any separation. No separation means that scores of signals were collected simultaneously. Therefore the signal assignment will be very intricate and even rather ambiguous, especially for tremendously complex samples like human serum and urine where thousands of metabolites may coexist.

The quantitative analysis linear ranges for the analytes varied from 100 to 1200 μM. The MS response tended to be flat when the concentration was over the linear ranges. The LODs of the analytes varied from 0.23 μM to 9.25 μM. G and A had the highest LOD. In Figure 2.11, G and A have the highest pI values among the 11 analytes separated in cation exchange mode. They were the last two analytes eluted in cation exchange mode. The higher the pI is, the more positively charged the analyte is, then the stronger its binding with the negatively charged stationary phase is; as a result, a higher concentration of H^+ is required to elute the analyte. If a higher TFA concentration in buffer B is to be applied, the recovery of G and A may be improved and lower LODs of A and G may be achieved. But one of the drawbacks is that a higher TFA concentration may hurt the subsequent mixed-mode separation.

The multidimensional separation method was also used to separate the components of the urine sample donated by a healthy adult. Figure 2.13 is the total ion chromatogram (TIC) that shows that total ions decreased from the cation exchange separation to the mixed-mode separation, to reversed phase separation, since most of the compounds are small and polar. Three urinanalysis runs with the same injection were performed on 3 consecutive days. Both the retention time and the characteristic m/z of the standards were used to extract the base peak chromatogram to quantitate the 24 analytes. The average calculated concentrations and CVs of the 24 analytes in the donated urine sample are listed in Table 2.2.

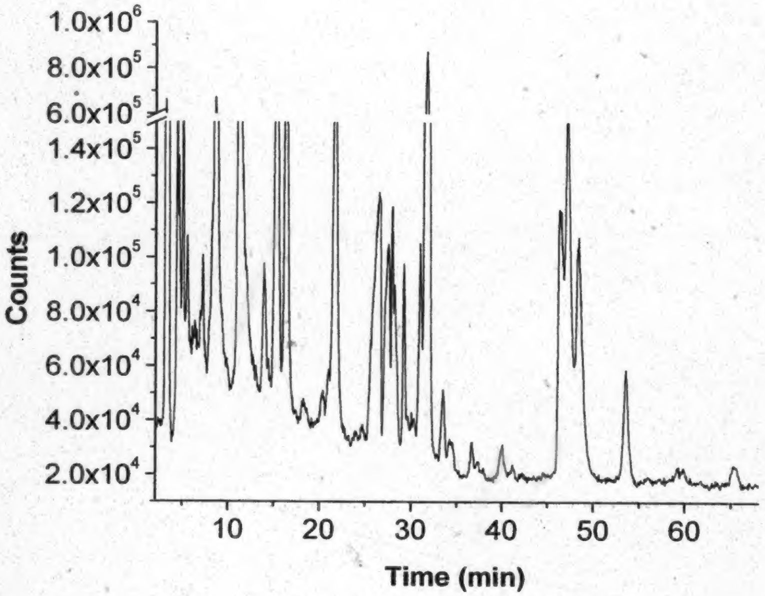


Figure 2.13. The TIC of a urine sample separated with the multidimensional separation method on one single mixed-mode HPLC column. HPLC conditions are shown in Figure 2.8.

2.4. Conclusions

The bifunctional HPLC column has been demonstrated in the literature to separate analytes that are too close in physical or chemical characteristics to be separated by a conventional column based on a single mode separation. Our work shows that the bifunctional column can provide both single-mode and mixed-mode separation mechanisms. This versatility provides benefits that may save the cost on column and overcome solvent compatibility problems that may be present between different types of columns during method development. This work shows that the method development with a bifunctional column should start with single-mode separation mechanisms instead of mixed-mode separation mechanisms, unless a single-mode separation mechanism cannot fulfill the objective because the complexity of mixed-mode HPLC method development is much higher than that of single-mode HPLC method development.

Ion-pair reagents have been used to form ion-pairs with polar compounds, such as amino acids, so that they can have usable retention times on reversed phase HPLC columns. However, this work suggests that the ion-pair reagents can increase retention of polar molecules on the mixed-mode column only when the ion-pair reagents are at a concentration greater than a critical concentration. If the concentration of an ion-pair reagent is below this critical concentration, the ion-pair reagent may actually reduce the analytes' retention on the reversed phase HPLC column. This work shows that the retention enhancement caused by ion-pair reagents varies widely for different analytes, which could lead to complexity in mixture separation; it may improve some compounds' separation factors, while it may also reduce other compounds' separation. It is also demonstrated that the retention enhancement caused by the ion-pair reagents is proportionally related to the pI of the analytes, the concentration and length of the ion-pair reagents.

We have developed a method of performing multidimensional HPLC separation on a single HPLC column. It is illustrated that the bifunctional column can be run in cation exchange mode, mixed-mode, and then reversed phase mode in series to form a multidimensional separation on a single column. A post-column TFA-fix is used to enhance the method's sensitivity, when it is combined with ESI MS. Different TFA-fixes have been evaluated and a possible mechanism is provided to explain their performance. In general, the sensitivity can be improved by lowering surface tension of the ESI droplets and replacing TFA to allow the formation of weaker ion-pairs with the analytes. Our method is then used to separate whole sets of amino acids and some of their metabolites. To our knowledge, this is the best separation method for amino acids ever achieved with MS-friendly HPLC conditions. Compared with regular on-line multidimensional HPLC, multidimensional HPLC on a single HPLC column can lower the cost significantly by eliminating the need for multiplex HPLC systems, columns and on-line traps.

Urinanalysis with a sample donated by a healthy adult is performed with the multidimensional separation method. Both the retention times and the characteristic m/z values of the standards are used to extract the base peak chromatogram. The peak intensities of the base peak chromatogram are used for quantitation. In the future a MS/MS detection mode with multiple reactions monitoring (MRM) may be developed to replace the MS detection mode to further improve detection sensitivity and confidence of compound identification.

2.5. References

1. Consden, R.; Gordon, A. H.; Martin, A.J.P. *Biochem. J.*, **1944**, 38, 224-232.
2. Liu, Z. Y.; Patterson, D. G.; Lee, M. L.,Jr.; *Anal. Chem.*, **1995**, 67, (21), 3840-3845.
3. Arey, J. S.; Nelson, R. K.; Xu, L.; Reddy, C. M. *Anal. Chem.*, **2005**, **77,** (22), 7172-7182.
4. Cai, H.; Stearns, S. D. *Anal. Chem.*, **2004**, 76, (20), 6064-6076.
5. Focant, J.-F.; Sjodin, A.; Turner, W. E.; Patterson, D. G., Jr. *Anal. Chem.*, **2004**, 76, (21), 6313-6320.
6. Hyotylainen, T.; Kallio, M.; Hartonen, K.; Jussila, M.; Palonen, S.; Riekkola, M.-L. *Anal. Chem.*, **2002**, 74, (17), 4441-4446.
7. Davies, I. L.; Bartle, K. D.; Williams, P. T.; Andrews, G. E. *Anal. Chem.*, **1988, 60,** (3), 204-209.
8. Libardoni, M.; Waite, J. H.; Sacks, R. *Anal. Chem.*, **2005**, 77, (9), 2786-2794.
9. Lu, X.; Cai, J.; Kong, H.; Wu, M.; Hua, R.; Zhao, M.; Liu, J.; Xu, G. *Anal. Chem.*, **2003**, 75, (17), 4441-4451.
10. Seeley, J. V.; Kramp, F.; Hicks, C. J. *Anal. Chem.*, **2000**, 72, (18), 4346-4352.
11. Shellie, R.; Marriott, P. J. *Anal. Chem.*, **2002**, 74, (20), 5426-5430.
12. Tang, K.; Li, F.; Shvartsburg, A. A.; Strittmatter, E. F.; Smith, R. D. *Anal. Chem.*, **2005**, **77,** (19), 6381-6388.
13. Moore, A. W.; Jorgenson, J. W., Jr. *Anal. Chem.*, **1995,** 67, (19), 3456-3463.
14. Yamabe, T.. *J. Chromatogr.* , **1973**, 83, 59-65.
15. Hancock, W. S.; Sparrow, J. T. *J. Chromatogr.*, **1981**, 206, (1), 71-82.

16. Lloyd, J. R.; Cotter, M. L.; Ohori, D.; Oyler, A. R. *Anal. Chem.*, **1987**, 59, (20), 2533-2534.

17. Saari-Nordhaus, R.; Anderson, J. *Anal. Chem.*, **1992,** 64, (19), 2283-2287.

18. Sun, L.; Carr, P. W. *Anal. Chem.*, **1995**, 67, (15), 2517-2523.

19. Huang, P.; Jin, X.; Chen, Y.; Srinivasan, J. R.; Lubman, D. M. *Anal. Chem.*, **1999**, 71, (9), 1786-1791.

20. Fu, H.; Xie, C.; Xiao, H.; Dong, J.; Hu J.; Zou H. *J. Chromatogr. A*, **2004**, 1044, (1), 237-244.

21. Nogueira, R.; Lämmerhofer, M.; Lindner, W. *J. Chromatogr. A*, **2005,** 1089, (1), 158-169.

22. Ferguson, P. L.; Iden, C. R.; Brownawell, B. J. *J. Chromatogr. A*, **2001**, 938, (1), 79-91.

23. Venkatramani, C. J.; Zelechonok, Y. *J. Chromatogr. A*, **2005**, 1066, (1), 47-53.

24. Washburn, M. P.; Ulaszek, R. R.; Yates, J. R., III *Anal. Chem.*, **2003**, 75, (19), 5054-5061.

25. Levy, A. L.; Chung, D. *Anal. Chem.*, **1953**, 25, (3), 396-399.

26. Hamerman, D.; Bartz, K. W.; Reife, A. *Anal. Chem.*, **1955**, 27, (10), 1524-1525.

27. Watanabe, K.; Nishiyama, M. Anal. Biochem. **1995**, 227, (1), 195-200.

28. Tuzimski, T.; Soczewiski, E. *J. Chromatogr. A*, **2002**, 961, (2), 277-283.

29. Horvath, C. G.; Lipsky, S. R. *Anal. Chem.*, **1967**, 39, (14), 1893-1893.

30. Mirza, U. A.; Chait, B. T. *Anal. Chem.*, **1994**, 66, (18), 2898-2904.

31. Banks, J. F.; Gulcicek, E. E. *Anal. Chem.*, **1997**, 69, (19), 3973-3978.

32. Qu, J.; Wang, Y.; Luo, G.; Wu, Z.; Yang, C. *Anal. Chem.*, **2002**, 74, (9), 2034-2040.

33. Karger, B. L.; Wong, W. S.; Viavattene, R. L.; Lepage, J. N.; Davies, G. *J. Chromatogr.*, **1978**, 167, 253-272.

34. Levin, S.; Grushka, E. *Anal. Chem.*, **1985**. 57, (9), 1830-1835.

35. Yukimoto, T.; Yoneda, H. *J. Chromatogr.*, **1981**, 210, (3), 477-485.

36. Madden, J. E.; Haddad, P. R. *J. Chromatogr. A*, **1998**, 829, (1), 65-80.

37. Rundlett, K. L.; Armstrong, D. W. *Anal. Chem.*, **1996**. 68, (19), 3493-3497.

38. Apffel, A.; Fischer, S.; Goldberg, G.; Goodley, P. C.; Kuhlmann, F. E.; *J. Chromatogr. A*, **1995**, 712, (1), 177-190.

39. Parvy, P.; Bardet, J.; Rabier, D.; Kamoun, P. C*lin. Chim. Acta*, **1995**, 235, (1), 1-10.

40. Schwarz, E. L.; Roberts, W. L.; Pasquali, M. *Clin. Chim. Acta*, **2005**, 354, (1), 83-90.

41. Jones, M. G.; Cooper, E.; Amjad, S.; Goodwin, C. S.; Barron, J. L.; Chalmers, R. A. *Clin. Chim. Acta*, **2005**, 361, (1), 150-158.

42. Clarke, A. P.; Jandik, P.; Rocklin, R. D.; Liu, Y.; Avdalovic, N. *Anal. Chem.*, **1999**, 71, (14), 2774-2781.

43. Ding, Y.; Yu, H.; Mou, S. *J. Chromatogr. A*, **2002**, 982, (2), 237-244.

44. Reddy, M. M. K.; Ghosh, P.; Rasool, S. N.; Sarin, R. K.; Sashidhar, R. B. *J. Chromatogr. A*, **2005**, 1088, (1), 158-168.

45. Nord, L. I.; Vaag, P.; Duus, J. O. Anal. Chem., **2004**, **76,** (16), 4790-4798.

Chapter 3

Rapid Bacterial Identification by LC-ESI MS/MS Analysis of Tryptic Peptides using a Monolithic Capillary Column

3.1. Introduction

Rapid detection and identification of bacteria plays a crucial role in medical microbiology and detection of biohazards.[1] Spectroscopy[2] and mass spectrometry (MS)[1, 3-7] have been extensively studied in microorganism detection through protein identification. However, only MS has the ability to perform bacterial spore identification[2] and fingerprinting identification[8]. In recent decades, MS has become an increasingly important technology platform for bacterial identification. The initial MS-based bacterial identification was primarily based on the comparison of the mass spectra of the proteins extracted from unknown bacteria to those from known bacteria. MALDI-MS[3, 5, 6] and ESI-MS[7] were used to measure masses of proteins extracted from microorganisms. The mass profiles were then searched against protein databases and the bacteria were identified based on the taxonomy of the identified proteins in the databases. This approach has the advantage of fast sample preparation. However, assignments to proteins based solely on mass values are limited by low mass accuracy, insufficient mass resolution, and lack of reproducibility of spectra. This approach often leads to ambiguous or even false identification. The other approach for characterization of microorganisms is similar to the "shotgun" approach used for many proteomic applications. Peptides derived from proteome-wide, enzymatic digestion of bacterial proteins are analyzed by high performance liquid chromatography (HPLC) electrospray ionization tandem mass spectrometry (ESI MS/MS).[1] Typically, for fast identification of bacteria, no more than one dimension of separation is applied. However, on a conventional packed particle column, complicated peptide mixtures from digestion of bacterial proteins cannot be efficiently separated with only a one-dimensional separation.

Speed of bacterial identification is of particular importance in many applications, including responses to bioterrorism attacks. Some novel concepts have been explored to shorten the analysis time. For example, specific proteins have been used as biomarkers and their identification, instead of global protein identification, can result in fast microorganism identification.[9, 10] Recently another innovative way to reduce the time per analysis is through MALDI ion trap/TOF tandem MS analysis of on-probe enzymatic digestion of proteins from microorganisms.[4] With the shot-gun approach, the major step limiting the speed of analysis is the separation. A reversed-phase separation using a traditional particle packed column usually takes 45-120 minutes.

Monolithic columns, a new generation of separation matrix, have attracted much attention[11-13] since their introduction[14-16]. They have been widely used in metabolite analysis of plasma,[17-20] bacteria,[21] and plant,[22] DNA,[23] and proteomics [24-27]. A monolithic column is made of a single piece of porous solid with large through channels (macro pores) for convective flow and high connectivity among mesopores on the skeleton for chromatographic interaction. A large number of inter-connected mesopores on the skeleton provide a large surface area to improve the retention, while the large macro pores allow the use of high mass transfer with a high flow-rate to achieve a fast separation. Thus the total separation time can be shortened significantly. Compared to conventional packed column, monolithic columns can achieve higher separation efficiency in much shorter elution time.

In this work, we report a proteomic approach for fast and reliable bacterial identification. The core of this method is the use of a capillary reversed phase monolithic column to separate bacterial tryptic peptides. Due to monolithic columns' high separation efficiency and high permeability, gradient time can be shortened to 2.5 min, compared to 45-120 min with conventional packed columns. Different gradient slopes and MS/MS scan speeds were investigated in detail to obtain the maximum number of identified proteins. The reproducibility and sensitivity were also examined. Finally, identification of four pure bacterial cultures as well as two bacterial mixtures was demonstrated.

3.2. Experimental

3.2.1. Chemicals and Reagents

Acetonitrile (ACN) (HPLC gradient grade) was obtained from Fisher (Fair Lawn, NJ). Water was purified by a Milli-Q UV plus Ultrapure system (Millipore, Mississauga, ON). Ammonium bicarbonate, dithiothreitol (DTT), and trifluoroacetic acid (TFA) were purchased from Sigma (St. Louis, MO), and trypsin (sequencing grade modified) from Promega (Madison, WI).

3.2.2. Bacteria Samples

Escherichia coli K-12 (*E. coli*, ATCC 47076), *Bacillus cereus* (ATCC 14579), *Lactococcus lactis* (ATCC 11454) and *Agrobacterium tumefaciens* (ATCC 33970) were ordered from the American Type Culture Collection. Bacterial cells were incubated under ATCC recommended conditions, harvested, washed with water, lyophilized, and stored at –80 °C.

3.2.3. Cell Lysis, Protein Digestion, and Sample Cleanup

1 mg *E. coli* cells were lysed with a French Press mini cell under 900 psi twice in 0.1 mM ammonium bicarbonate, 1 mM DTT buffer at 4 °C. The suspension was centrifuged at 11750 g. 0.2 M ammonium bicarbonate was added to the protein extract to adjust the pH to 8.3, then the proteins were denatured by incubating at 95 °C for 10 min, and digested by trypsin at a ratio of 1:50 (w/w) at 37 °C overnight. TFA was used to quench the digestion reaction. A Ziptip C18 (Millipore, Mississauga, ON) was used to desalt the digested peptide mixture before HPLC analysis. *Bacillus cereus*, *Lactococcus lactis* and *Agrobacterium tumefaciens* cells were lysed in 0.2 M ammonium bicarbonate by sonication (Branson probe sonicator, Branson Ultrasonics Corp., Banbury, CT) for 1 min (1 pulse/s with 0.75 s pulse duration), respectively. The suspension was centrifuged at 11750 g, and then the proteins in the supernatant were denatured by incubating at 95 °C for 5 min, and digested by trypsin at a ratio of 1:20 (w/w) at 37 °C for 10 min. All the bacteria were treated in biosafety (I) lab.

3.2.4. HPLC and Electrospray Ionization Mass Spectrometry

An Agilent 1100 HPLC system with binary pump was used for gradient generation and solvent delivery. A monolithic (poly(styrenedivinylbenzene) copolymer) capillary column 5 cm × 200 μm i.d. from LC packings (Amsterdam, The Netherlands) was used for HPLC-ESI-MS (or MS/MS). Mobile phase A contained 0.05% TFA in water, B contained 0.04% TFA in ACN. The gradient profiles were set as follows: B% was changed from 5% to 10% in 1 minute, from 10% to 35% in 2.5, 5, 10 and 15 minutes, respectively, from 35% to 90% in 1 minute and held at 90% for 1 minute, and finally changed back to 5% in 1 minute. Each HPLC run consisted of a 2 μL injection of peptide mixture resulting from the digestion of approximately 1 μg of bacterial proteins.

The HPLC effluents were analyzed by an LCQ Advantage/LCQ Deca ion trap mass spectrometer controlled by Xcalibur (v 1.3) (ThermoFinnigan, San Jose, CA). The LCQ was set to acquire a full-scan mass spectrum between *m/z* 400 to 1400 followed by 3 to 6 data-dependent product ion mass spectral scans depending on microscan number. The collision energy was set at 35%. Other parameters for Xcalibur are set as a duration of 0.5 min with a 1 min exclusion window.

3.2.5. Data Processing

The product ion mass spectra were searched against two databases, a non-redundancy database and *E. coli* protein database with TurboSEQUEST (ThermoFinnigan). Both databases were downloaded from the National Center for Biotechnology Information (NCBI) site in a FASTA format. Three parameters, Xcorr, ΔCn and peptide hits were used to rank the significance of product ion mass spectral matching with the peptide sequence in the database. For singly charged ions Xcorr was at least 1.9, doubly charged ions Xcorr was at least 2.2, for triply charged ions Xcorr was at least 3.75. SEQUEST search results were filtered with specifications of ΔCn greater than 0.1 and the first peptide hits greater than 0. Only the unique peptides were used for statistical analysis and comparison.

MASCOT (Matrix Science Ltd, 8 Wyndham Place, London W1H 1PP, UK) was also used as a search engine to search against the NCBI bacterial database for bacterial mixture identification. It was found that MASCOT worked better than SEQUEST in terms of spectral matching properties, and thus it was used throughout the bacterial identification work. Only the first rank peptides were considered as identified peptides. Only proteins with one peptide scoring higher than 45 or at least two peptides scoring higher than 30 were considered as identified proteins. All of the peptides were assigned once only. If a peptide could be assigned to more than one protein, the peptide would be assigned to the protein with the highest scores.

3.3. Results and Discussion

3.3.1. Fast Gradient

Due to the ion suppression effect in LC-ESI MS/MS analysis, high quality MS/MS data generation depends on the resolution of chemical species in LC. However, in proteomics applications, the complexity of the protein or peptide mixture overwhelms the capability of any existing separation technology. Multidimensional separations can achieve higher separation efficiency, but require much longer run time and much more sample. For fast bacterial identification, one-dimensional separation is used. Monolithic columns are ideal candidates to solve the dilemma of achieving both high separation speed and efficiency. The concept of monolithic media for HPLC columns is relatively new.[33-35] One kind of monolithic column is made by the polymerization of an organic monomer in the presence of porogens.[33-35] Different types and concentration ratios of monomer and porogens yield different mesopores with size ranging from 0.5 μm to 2 μm, providing efficient separation and large macropores to enable faster mass transfer. This unique structure allows the separation to be made with low flow impedance and high resolution at high flow rates and fast gradients.

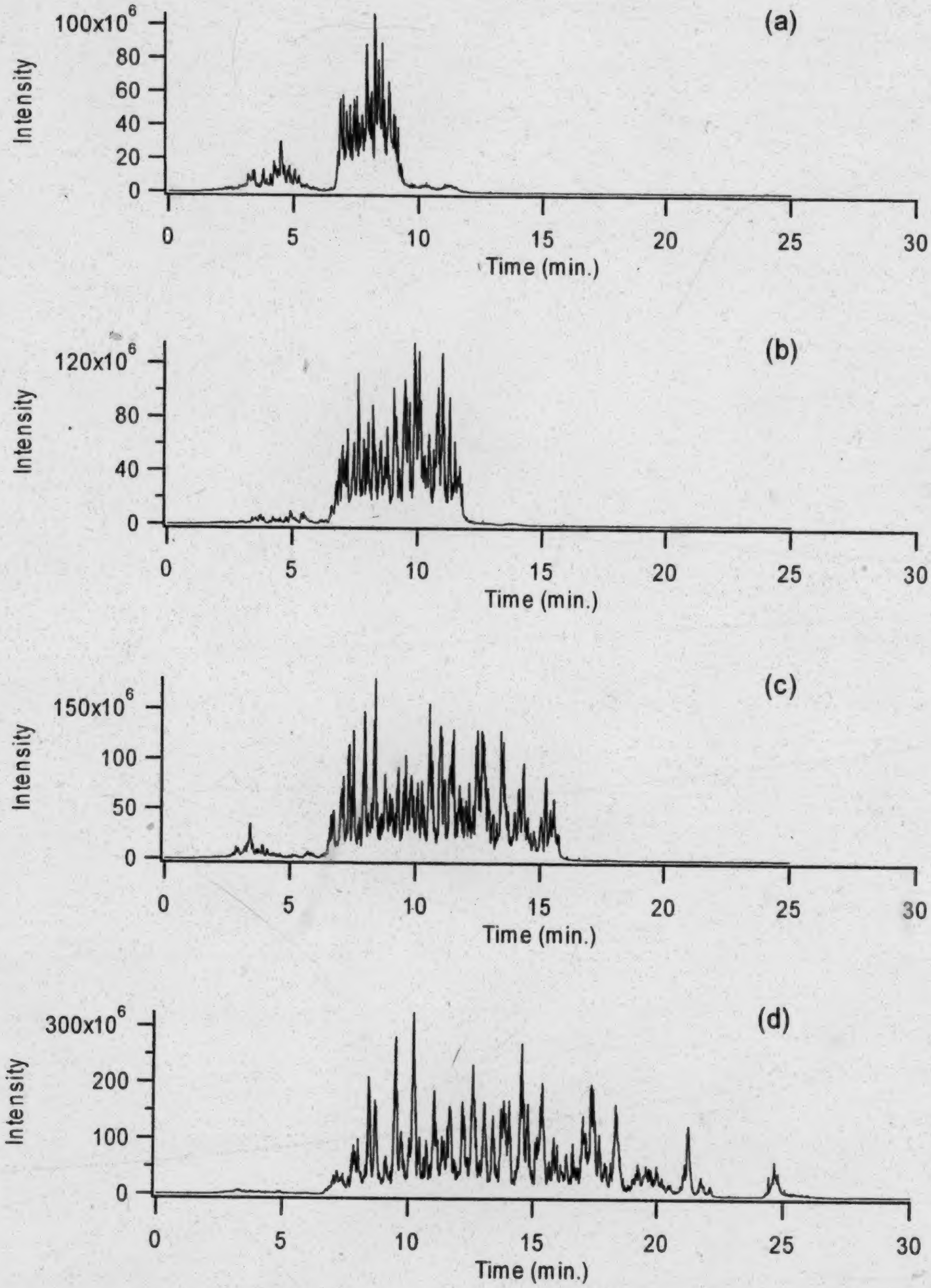


Figure 3.1. Base peak intensity chromatograms of capillary monolithic RP HPLC-ESI MS with different gradient times from 10% B to 35% B; (a) 2.5 min, (b) 5 min, (c) 10 min, (d) 15 min, respectively. Flow rate is 2 μL/min.

Figure 3.1 shows the base peak ion chromatograms from the separations of the tryptic digest of an *E. coli* cell extract using the monolithic column. In order to optimize the chromatographic performance of the monolithic column, TFA was chosen as the ion-pairing reagent to sharpen peak shapes, despite its ion suppression effect.[31] A low concentration of TFA was used to minimize the ion suppression effect. Figure 3.1 shows that all peptides were eluted within a time window that closely matched the time for the gradient from 10% B to 35% B. For the following discussion, the time to deliver this gradient window is termed as the peptide separation window (PSW). It is obvious that with longer PSW, more peptides can be resolved and more MS/MS data can be obtained, therefore more proteins can be identified. For fast bacterial identification, a PSW of as short as 2.5 min can be used. More than 80 peptides were identified within a 2.5 min PSW, which is sufficient for bacterial identification. With traditional packed columns, this separation can not be easily achieved within such short periods using conventional column HPLC systems. It is also noted that the base peak intensity increases as the PSW increases. This is, at least partially, due to the better peptide recovery rendered by the increased elution time; hence, higher base peak intensity.

3.3.2. MS and MS/MS scan speed

Although monolithic columns can improve separation efficiency dramatically compared to traditional column packing, peptide co-elution across each chromatographic peak is unavoidable due to the complex nature of the peptide mixture. MS and MS/MS scan speed were optimized to be compatible with fast LC separation, thus maximizing the number of identified proteins.

In data-dependent MS/MS acquisition, one cycle of data acquisition consists of one survey MS scan followed by a selected number of MS/MS scans. One MS scan including AGC prescan and a number of microscans. And one MS/MS scan including AGC prescan, ion isolation, ion activation, and a number of microscans. To optimize MS and MS/MS scan speed, the only adjustable parameter is the number of microscans, while other parameters are either fixed or determined by the instrument automatically. The results of microscans were averaged for each MS and MS/MS spectrum to achieve better S/N, resulting in high quality MS and MS/MS spectra and a high success rate for peptide matching in database search. Increasing the number of microscans reduces the duty cycle of the instrument. Thus the frequency of data acquisition is reduced, resulting in fewer MS/MS spectra over the chromatographic peak.

To maximize the number of identified peptides, another adjustable parameter is the number of MS/MS scans following each survey MS scan. At every fixed microscan number (M), altering the number of MS/MS scans (N) to obtain the maximum MS/MS spectra across chromatography peak. One cycle of data acquisition time, determined by M and N, is set up to be the same as the half peak width. In the following discussion, (M, N) will be used to denote the different combinations of the number of microscans and the number of analytical MS/MS scans in the instrument set-up, respectively. Due to the very high separation efficiency of monolithic column, all peaks showed up very sharply when a fast gradient was applied. The capillary monolithic column gave an average half peak width of 4 s in a 5-min gradient for peptides. An estimate based on the scan speed of 5,500 amu/s (the normal scan speed for LCQ series ion tap MS) indicates that *(1,6), (2,4)* and *(3,3)* all result in a duty cycle of approximately 4 s. The quality of the MS/MS spectra were then evaluated by the number of matched peptides and identified proteins from a SEQUEST search against the *E. coli* database and the results are summarized in Figure 3.2 (a).

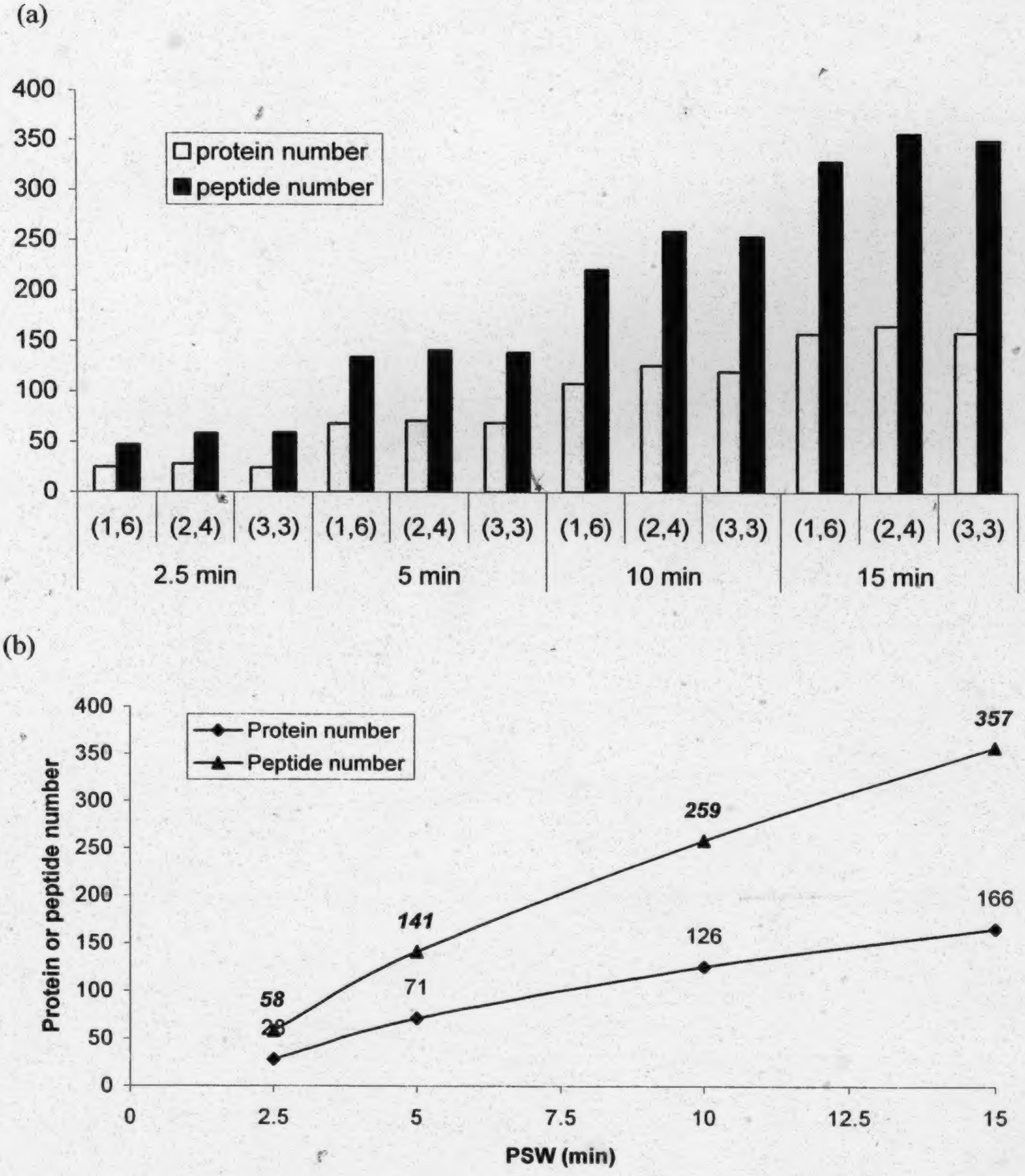


Figure 3.2 (a) The average numbers of matched peptides and identified proteins obtained from 3 replicate experiments with different PSWs and MS instrument settings. (b) The number of matched peptides and identified proteins from each (2,4) setting plotted against PSW.

Figure 3.2 clearly shows that slower gradients gave more identified peptides and proteins. From a separation point of view, a slower gradient resulted in better peptide separation, the ion suppression was reduced and then more peptides could be ionized and potentially identified through MS/MS. Concerning the MS part, a slower gradient would potentially allow each LC peak to have a longer MS/MS response window, and therefore would potentially generate more MS and MS/MS spectra. For every PSW setting, the (2, 4) combination generated the most identified proteins and peptides. With a (1, 6) setting not all ions could be swept into the detector and this caused the missed detection of some ions and decreased the number of identified peptides and proteins. If the microscan number is greater than one, spectra from the continuous microscans will be averaged, and the chance of missing in detection will decrease. Obviously, when more than one microscan was applied in the scan event, spectra with better quality could be obtained, which could provide easier assignment during peptide MS/MS database search. This is the reason why one microscan gave the most MS/MS spectra (data not shown), but had actually the least identified peptides and proteins. By contrast, when a (3, 3) setting was applied, although the spectra with the best quality were attained, less proteins and peptides were identified compared with the results from two microscans, due to lower acquisition frequency. Consequently, the optimized scan event pattern for the fast gradient was (2, 4).

It is also interesting to see from Figure 3.2 (b) that, when PSW increased from 10 min to 15 min, the number of identified peptides increased much faster than the identified proteins. Actually, when PSW was longer than 15 min, although the number of the identified peptides still increased, the protein number reached a plateau (data not shown). This indicates that, when PSW was longer than 10 min, most of the extra-identified peptides were from already identified proteins. For the purpose of fast bacteria identification, to identify the most proteins in shortest time, 10 min PSW will be the most efficient LC setting. Nonetheless, the numbers of peptides and proteins generated even within a 2.5 min PSW should be sufficient for a positive bacterial identification.

These results demonstrate that, in order to perform prolific and efficient protein identification using a "shotgun" approach by LC-ESI MS/MS, LC gradient, spectral acquisition speed, and spectral quality have to be considered as integrated elements of the overall experimental design.

3.3.3. Reproducibility

In order to achieve a reliable bacterial identification, reproducible protein identification is a prerequisite. The reproducibility of protein identification in this study was evaluated by repeating the experiment 3 times under the same experimental conditions. The percentage of the number of proteins detected in each run over total detected proteins in all three runs was used as a measure of reproducibility.

The results shown in Figure 3.3 indicate that for all PSWs the reproducibility increases as the number of microscans increases. The reason is that, when more than one microscan was applied, by averaging spectra from the microscans, more ions would be detected, leading to better reproducibility of MS spectra. It needs to be pointed out that 2 microscans gave almost the same reproducibility as 3 microscans. This means that, in terms of reproducibility, the scan event pattern with 2 microscans is satisfactory in this experiment. Moreover, Figure 3.3 shows that 2 microscans gave the most proteins identified in total. This result matches the result shown in Figure 2 that 2 microscans gave the highest number of identified proteins, and further confirms that the scan event pattern with 2 microscans is the optimized one. Under the same (M, N) setting, the longer the gradient, the better the reproducibility would be. Due to the complexity of the analyte, there were many peptides eluted at the same time. These peptides competed with each other in the ESI process, which resulted in ion suppression and decrease in reproducibility. A slower gradient can resolve the problem to some degree, but slower gradients mean longer analysis times and are not favorable for fast protein identification.

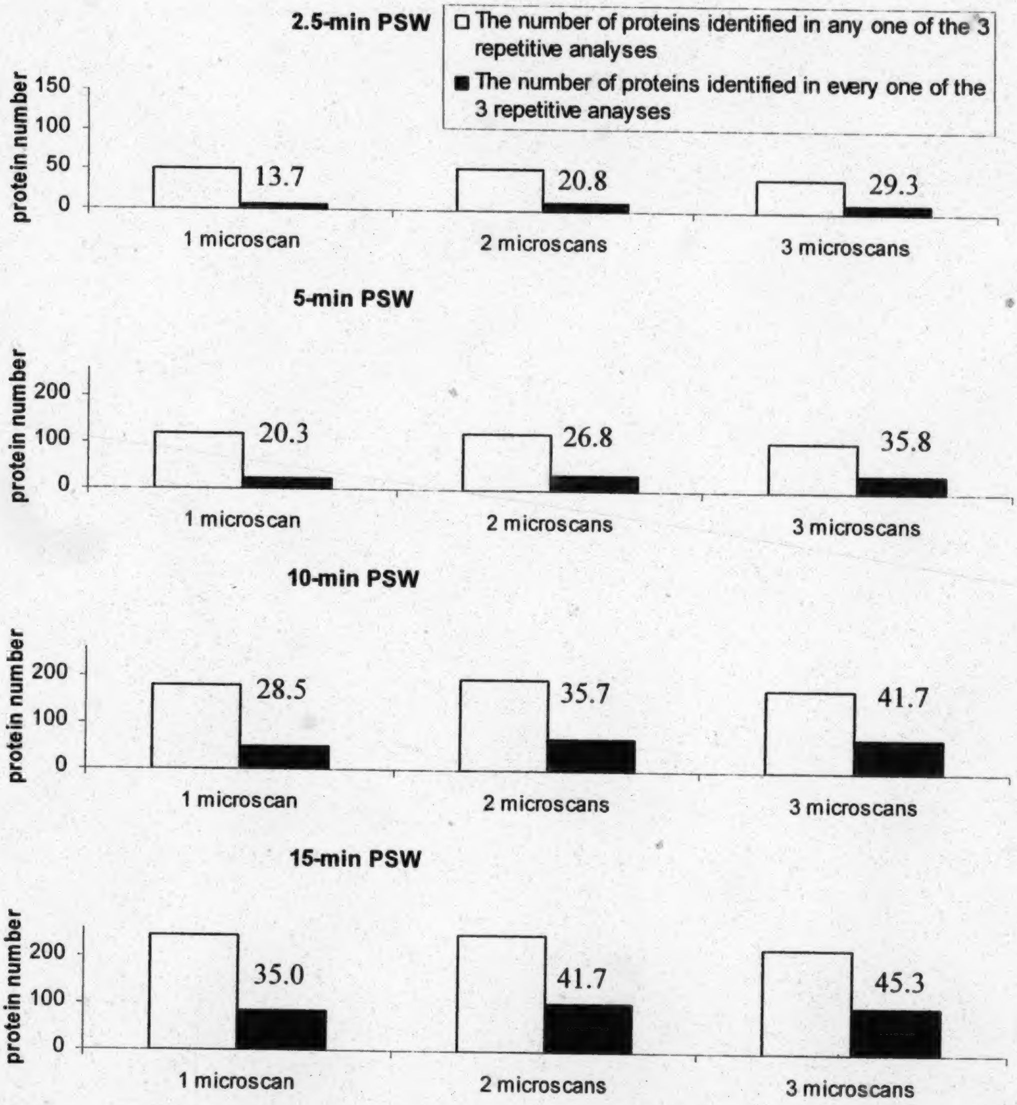


Figure 3.3. Reproducibility analysis from 3 replicate experiments with different PSWs. Flow rate was 2.0 μL/min and 1 μg of the digest of *E. coli* was injected for each experiment.

Figure 3.4 shows a more detailed view of reproducibility among 3 runs with a 10 min gradient and 2 microscans. There were a total of 204 proteins identified in 3 replicate runs. Protein IDs assigned with a number from 1 to 204 are used in all three X-axis's in Figure 3.4. The Y-axis is the number of unique peptides detected from each protein. Figure 3.4(a), 3.4(b), and 3.4(c) show the results from the first, second and third run, respectively. In Figure 3.4(a) the first 55 proteins show decreasing peptide hits from 15 to 2 in order. A similar trend is also shown in Figure 3.4(b) and 3.4(c). In general, the confidence and reproducibility of protein identification increase as the number of peptides detected from the protein increases. The first 55 proteins in Figure 3.4(a) were identified with more than 2 peptide hits, and almost 90% of these proteins were also identified in 2 other runs. Table 3.1 summarizes the results of this study.

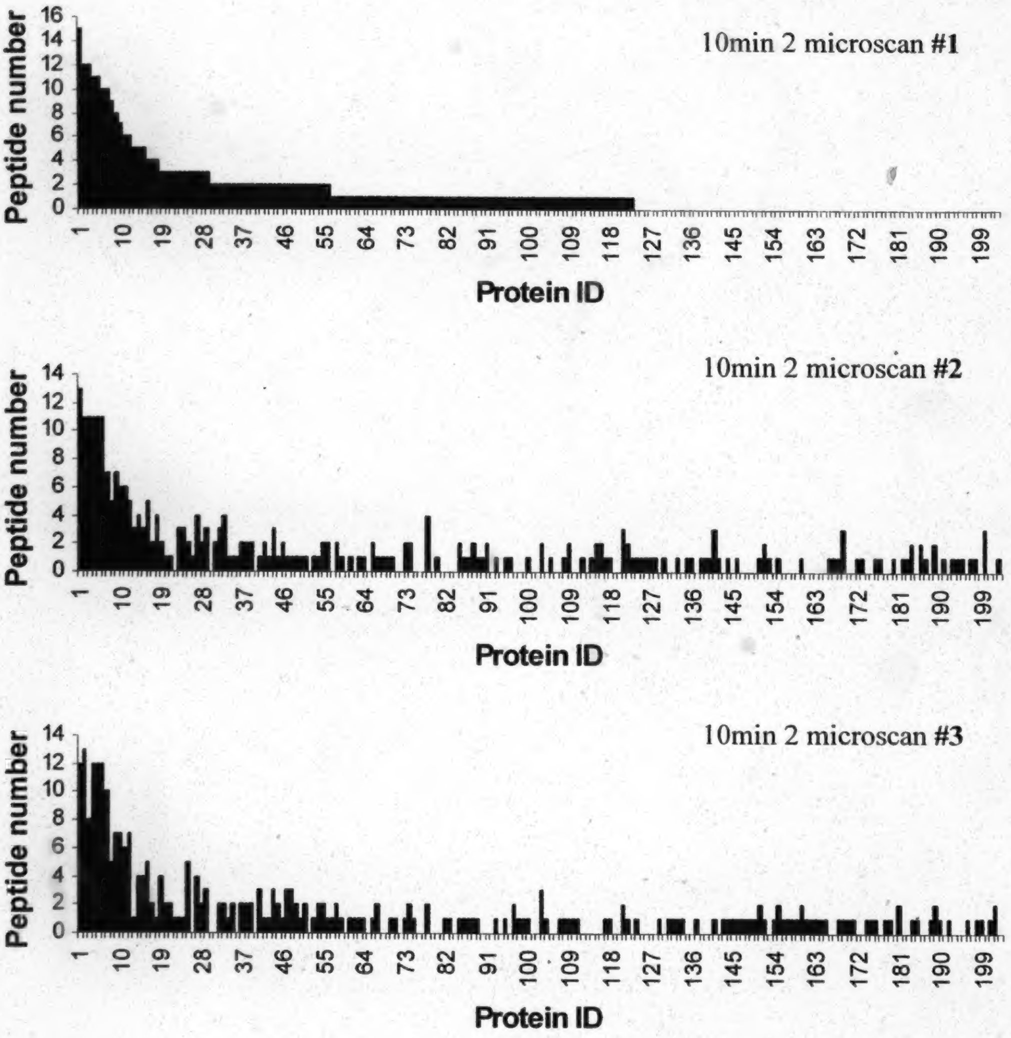


Figure 3.4. Zoomed in reproducibility of one-dimensional capillary monolithic RP HPLC-ESI-MS/MS with a 10-min HPLC gradient and scan event of 4 MS/MS and 2-microscans. The protein IDs in (a), (b), and (c) X-axis are completely identical to each other.

Table 3.1. A reproducibility breakdown from the replicate experiments using a 10 min PSW and *(2, 4)* instrument set-up.

	Protein in #1 run with at least 1 peptide detected	Protein in #2 run with at least 1 peptide detected	Protein in #3 run with at least 1 peptide detected	Averaged protein identity repeatability in %
The number of proteins identified in each run	116	126	135	
The number of proteins shown in every one of the 3 repeated runs	70	70	70	56%

	Protein in #1 run with at least 2 peptides detected	Protein in #2 run with at least 2 peptides detected	Protein in #3 run with at least 2 peptides detected	Averaged protein identity repeatability in %
The number of proteins identified in each run	45	47	41	
The number of proteins shown in every one of the 3 repeated runs	40	38	36	86%

Table 3.1. A reproducibility breakdown from the replicate experiments using a 10 min PSW and *(2, 4)* instrument set-up.

	Protein in #1 run with at least 3 peptides detected	Protein in #2 run with at least 3 peptides detected	Protein in #3 run with at least 3 peptides detected	Averaged protein identity repeatability in %
The number of proteins identified in each run	25	24	21	95%
The number of proteins shown in every one of the 3 repeated runs	23	22	21	

For the proteins identified by one or more unique peptide, on average only 56% of all proteins were identified in all three replicate runs. However, as this criterion becomes more stringent, the reproduciblity dramatically improves. For the proteins identified with 2 or more peptides, on average 86% of all proteins were identified in all three replicate runs. If the peptide hits threshold was set to 3, then 95% of these proteins on average could be identified in all three replicate runs. It is evident that the major variation in the "shotgun" approach for protein identification results from those proteins identified by a single peptide. Figure 3.5 shows the distribution of proteins identified by one, two, or three and more peptides. Up to 63% of the proteins were identified by only one peptide.

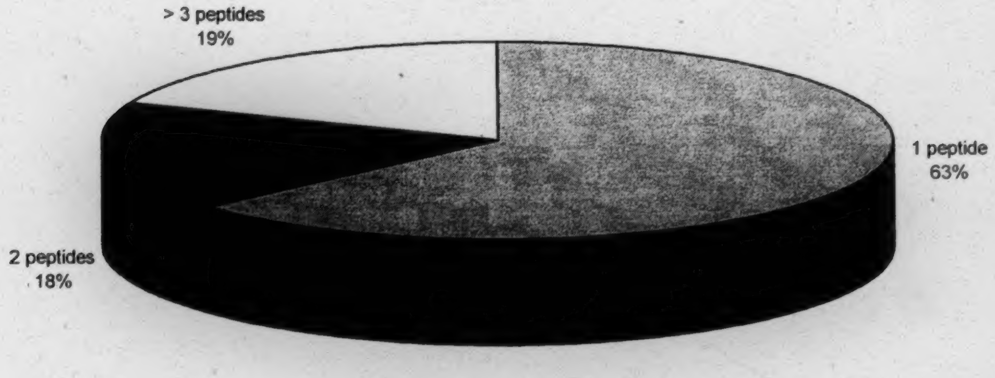


Figure 3.5. Distribution of the identified proteins using different numbers of peptides.

Generally, high abundance proteins have a better chance to be identified with multiple peptide hits. Unfortunately, concentrating low abundance proteins is difficult to accomplish in the one-dimensional separation required for fast identification. Although longer gradient times can help to improve separation and hence increase the number of peptides detected to a certain degree, they sacrifice the speed of identification. Moreover, for fast bacterial identification through peptide MS/MS database search, global protein information is not necessary. In fact, protein taxonomy profiles that can be classified to a specific microorganism are more important. Our work indicates that, for bacterial identification, as long as only the proteins are identified by two or more peptides, a satisfactory reproducibility of 90% or higher can be achieved from the monolithic LC ESI MS/MS analyses.

3.3.4. Sensitivity

Table 3.2 lists the number of peptides and proteins identified from a series of gradual dilution experiments that were made to gauge the detection sensitivity of the present technique.

Table 3.2. Overall sensitivity using two different PSWs. All experiments were performed on a LCQ Deca ion trap MS. Acetic acid was used as the ion-pairing reagent and the flow rate was set at 2.5 μl/min. There are around 1.0 X 10^9 cells in 1 mg *E. coli.* This number was obtained through counting the cells under microscope and was used to estimate the number of *E. coli* cells in samples for cell lysis.

PSW	Number of *E. coli* cells	Protein number	Peptide number
2.5 min	500,000	11	15
	400,000	10	12
	200,000	7	9
15 min	500,000	19	37
	400,000	13	26
	100,000	5	6
	80,000	1	1

Somewhat different experimental conditions from the last sections were employed. In the present work, 0.1% acetic acid instead of TFA was used to minimize ion suppression; a higher flow rate of 2.5 μL/min instead of 2 μL/min was used to match the fast gradient; and an LCQ Deca mass spectrometer instead LCQ Advantage was used for better instrument sensitivity. Using a 2.5 min PSW, an injection of a peptide mixture corresponding to 200,000 cells resulted in matching 9 peptides from which 7 proteins were identified. For a 15 min PSW, only 100,000 cells were needed to match 6 peptides and identify 5 proteins. The sensitivity was found to be instrument dependent. For the same experiments performed on the LCQ Advantage ion trap MS, an up to 20-fold decrease in the limit of detection (LOD) was observed. With the emergence of a new generation of mass spectrometers, such as linear ion trap mass spectrometers, and continuous improvement in instrument sensitivity, it is anticipated that protein identification from 10,000 cells or less can be implemented in the very near future. Further optimization in protein extraction, digestion, and sample injection will also help for identification of bacteria using a smaller number of cells.

3.3.5. Rapid Bacterial Identification by Monolithic LC-ESI MS/MS

Finally, the applicability of bacterial identification using monolithic LC-ESI MS/MS was evaluated by the analyses of four pure bacterial cultures and two bacterial mixtures. The MS/MS spectra were searched against the NCBI bacterial database by both SEQUEST and MASCOT. Simply, the percentage of identified proteins assigned to a specified organism over the total identified proteins assigned to all database organisms was used to score and identify the bacterium. It was found, however, that the same data set searched by SEQUEST generated a much lower score than that by MASCOT. This may be related to the difference in statistics algorithms used by the two software search engines.

Figure 3.6 shows the MASCOT search results for *E. coli* obtained using a *(2,4)* set-up and different PSWs. The scores range from 86% for a 2.5 min PSW to 93% for a 15 min PSW. No significant difference was found among all PSWs. A protein profile generated from as short as a 2.5 min PSW is sufficient for bacterial identification. It has to be pointed out that for a 2.5 min PSW, the total LC analysis time is actually 6.5 min (see Experimental), which is far shorter than the LC-MS/MS time required with a conventional packed column. To address the speed of analysis, a 2.5 min PSW with a slightly increased flow rate of 2.5 μL/min was used in all the following experiments.

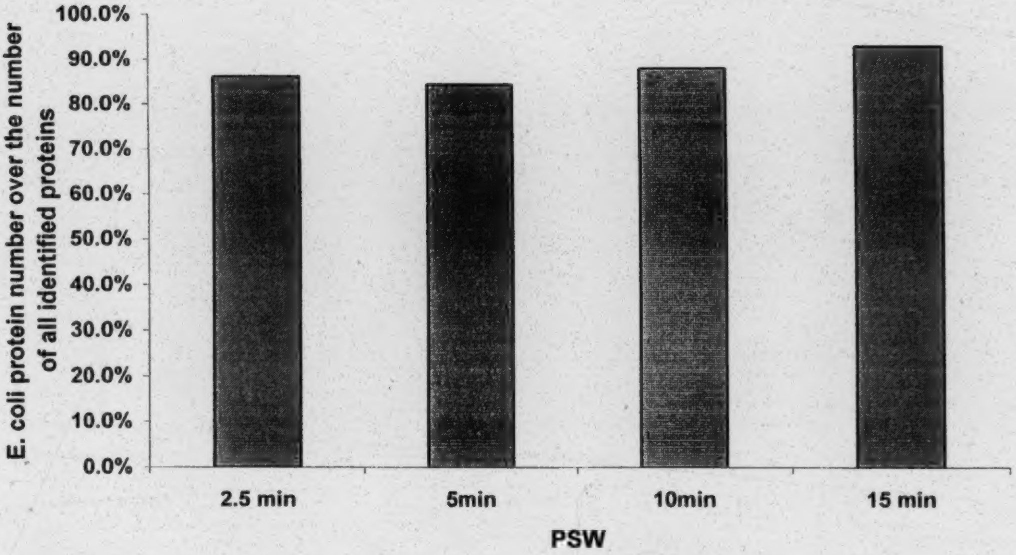


Figure 3.6. Bacterial identification score (%) from different PSWs. Flow rate was 2.0 μL/min and 1 μg of the digest of *E. coli* was injected for each experiment. The MS/MS spectra were searched against non-redundant bacterial database from NCBI using MASCOT.

Sonication and French Press are two widely used methods for protein extraction. The typical times required for sonication and French Press are about 3 min and 40 min, respectively. But the mild French Press gives higher extraction efficiency than sonication. To speed up the fast bacterial identification further, the French press was substituted for sonication with extracting proteins from *E. coli*, *Bacillus cereus, Lactococcus lactis,* and *Agrobacterium tumefaciens* in the following applications. The results of these analyses are summarized in Table 3.3. A score of 96% or higher was achieved for all four bacteria, showing high confidence in bacterial identification within a 10 min analysis time.

Table 3.3. The number of identified proteins from pure bacterial cultures and bacterial mixtures using a 2.5 min PSW. Flow rate was set at 2.5 μL/min. MS/MS data were searched against NCBI bacterial database using MASCOT.

Sample	Number of proteins Identified		Percent of target protein
	Target protein	Total protein	
E. coli	21	22	96%
Bacillus cereus	15	15	100%
Lactococcus lactis	29	30	97%
Agrobacterium tumefaciens	28	29	97%

In real world applications, such as water quality monitoring and detection of bio-warfare reagents on the battlefield, the target pathogenic microrganism is often present with interference from a number of background bacteria. In order to explore the specificity and relative sensitivity of this approach for bacterial identification using a monolithic column, two mixtures of the tryptic digests from A*grobacterium tumefaciens/Bacillus cereus/ Lactococcus lactis* (w/w/w, 3:2:1) and *Agrobacterium tumefacien/Lactococcus lactis* (w/w, 10:1) were analyzed. Figure 3.7 (a) shows results from the analysis of a mixture of peptides from, *Bacillus cereus, Agrobacterium tumefaciens* and *Lactococcus lactis* (w/w 3:2:1) with the developed method. All 31 identified proteins came from the three target bacteria, *Agrobacterium tumefaciens*, *Bacillus cereus*, and *Lactococcus lactis*. When the composition discrepancy became larger, as in the mixture of *Agrobacterium tumefaciens* and *Lactococcus lactis* (w/w, 10:1), some proteins were misidentified (Figure 3.7 (b)). The exact cause is hard to discern. Nonetheless, the proteins identified as from the target bacteria, *Agrobacterium tumefaciens* and *Lactococcus lactis*, still form the two largest demographic groups. The relative sensitivity of the method for bacterial mixtures, roughly equal to the weight ratio of 10:1, is comparable to in situ proteolytic digestion and analysis with qIT-TOF MS.[4]

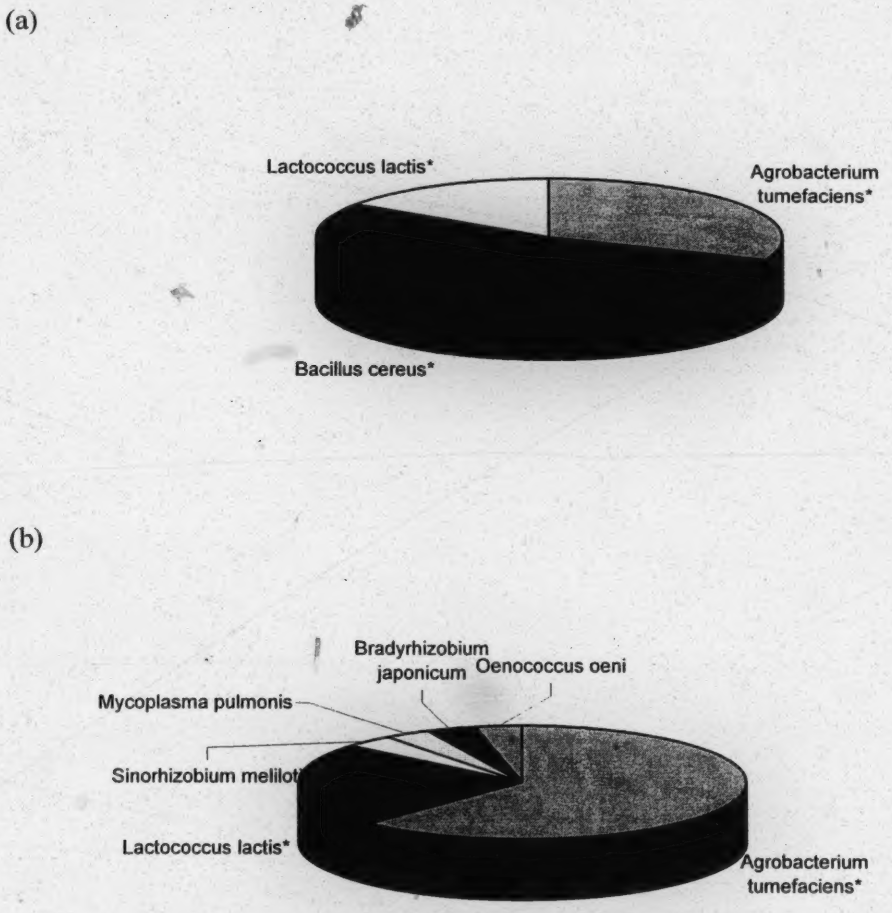


Figure 3.7. (a) Diagram showing the distribution of the identified proteins. 1 μg of the tryptic digest from *Agrobacterium tumefaciens, Bacillus cereus* and *Lactococcus lactis* (w/w/w, 3:2:1) was injected. (b) Diagram showing the distribution of the identified proteins. 1 μg of the tryptic digest from *Agrobacterium tumefaciens,* and *Lactococcus lactis* (w/w, 10:1) was injected. Flow rate was 2.5 μL/min and a *(2, 4)* instrument set-up was used for all experiments. The MS/MS spectra were searched against non-redundant bacterial database from NCBI using MASCOT.

3.4. Conclusions

A new method for fast identification of bacteria is described. The fast and efficient separation was achieved through one-dimensional HPLC with a capillary monolithic column. Benefiting from its unique biporous structure, the monolithic column demonstrated very high separation efficiency, capable of separating tryptic digests of bacterial proteins with a gradient of 2.5 min. It was demonstrated that a combination of 2 microscans for each MS survey scan, followed by 4 analytical MS/MS scans was the best compromise between data acquisition speed and spectral quality. Detection sensitivity was investigated and found to be dependent on the gradient speed. With the current setup and the use of a 2.5-min gradient, 200,000 cells were required for bacterial identification. Finally, this method was shown to be applicable for reliable identification of several pure bacterial cultures and simple mixtures of bacteria.

3.5. References

1. Dworzanski, J. P.; Snyder, A. P.; Chen, R.; Zhang, H.; Wishart, D.; Li, L. *Anal. Chem.* **2004,** 76, (21), 6492-6499.
2. Thompson, S. E.; Foster, N. S.; Johnson, T. J.; Valentine, N. B.; Amonette, J. E. *Appl. Spectrosc.* **2003,** *57*, 893 -899.
3. Pineda, F. J.; Antoine, M. D.; Demirev, P. A.; Feldman, A. B.; Jackman, J.; Longenecker, M.; Lin, J. S. *Anal. Chem.* **2003,** *75*, 3817-3822.
4. Warscheid, B.; Jackson, K.; Sutton, C.; Fenselau, C. *Anal. Chem.* **2003,** *75*, 5608-5617.
5. Ho, Y.-P.; Hsu, P.-H. *J. Chromatogr. A* **2002,** *976*, 103-111.
6. Wang, Z.; Dunlop, K.; Long, S. R.; Li, L. *Anal. Chem.* **2002,** *74*, 3174-3182.

7. Krishnamurthy, T.; Davis, M. T.; Stahl, D. C.; Lee, T. D. *Rapid Commun. Mass Spectrom.* **1999,** *13*, 39-49.

8. Jarvis, R. M.; Goodacre, R. *Anal. Chem.* **2004,** *76*, 40-47.

9. Cargile, B. J.; McLuckey, S. A.; Stephenson, J. L., Jr. *Anal. Chem.* **2001,** *73*, 1277-1285.

10. Vaidyanathan, S.; Winder, C. L.; Wade, S. C.; Kell, D. B.; Goodacre, R. *Rapid Commun. Mass Spectrom.* **2002,** *16*, 1276-1286.

11. Minakuchi, H.; Nakanishi, K.; Soga, N.; Ishizuka, N.; Tanaka, N. *Anal. Chem.* **1996,** *68*, 3498-3501.

12. Motokawa, M.; Kobayashi, H.; Ishizuka, N.; Minakuchi, H.; Nakanishi, K.; Jinnai, H.; Hosoya, K.; Ikegami, T.; Tanaka, N. *J. Chromatogr. A* **2002,** *961*, 53-63.

13. Vervoort, N.; Gzil, P.; Baron, G. V.; Desmet, G. *Anal. Chem.* **2003,** *75*, 843-850.

14. Hjerten, S.; Liao, J. L. *J. Chromatogr. A* **1988,** *457*, 165-174.

15. Hjerten, S.; Liao, J. L.; Zhang, R. *J. Chromatogr.* **1989,** *473*, 273-275.

16. Svec, F.; Frechet, J. M. J. *Anal. Chem.* **1992,** *64*, 820-822.

17. Deng, Y.; Wu, J.-T.; Lloyd, T. L.; Chi, C. L.; Olah, T. V.; Unger, S. E. *Rapid Commun. Mass Spectrom.* **2002,** *16*, 1116-1123.

18. Lutz, E. S. M.; Markling, M. E.; Masimirembwa, C. M. *J. Chromatogr. B* **2002,** *780*, 205-215.

19. Zeng, H.; Deng, Y.; Wu, J.-T. *J. Chromatogr. B* **2003,** *788*, 331-337.

20. Hsieh, Y.; Wang, G.; Wang, Y.; Chackalamannil, S.; Korfmacher, W. A. *Anal. Chem.* **2003,** *75*, 1812-1818.

21. Jia, L.; Liu, B.-F.; Terabe, S.; Nishioka, T. *Anal. Chem.* **2004,** *76*, 1419-1428.

22. Tolstikov, V. V.; Lommen, A.; Nakanishi, K.; Tanaka, N.; Fiehn, O. *Anal. Chem.* **2003,** *75*, 6737-6740.

23. Premstaller, A.; Oberacher, H.; Huber, C. G. *Anal. Chem.* **2000,** *72*, 4386-4393.

24. Walcher, W.; Timperio, A.-M.; Zolla, L.; Huber, C. G. *Anal. Chem.* **2003,** *75*, 6775.

25. Premstaller, A.; Oberacher, H.; Walcher, W.; Timperio, A. M.; Zolla, L.; Chervet, J.-P.; Cavusoglu, N.; van Dorsselaer, A.; Huber, C. G. *Anal. Chem.* **2001,** *73*, 2390-2396.

26. Barroso, B.; Lubda, D.; Bischoff, R. *J. Proteome Res.;* **2003,** *2*, 633-642.

27. Tanaka, N.; Kimura, H.; Tokuda, D.; Hosoya, K.; Ikegami, T.; Ishizuka, N.; Minakuchi, H.; Nakanishi, K.; Shintani, Y.; Furuno, M.; Cabrera, K. *Anal. Chem.* **2004,** *76*, 1273-1281.

28. D.J., P.; Capparella, M.; Neue, M.; Fallah, Z. E. *Journal of Pharm. Biomed. Anal.* **1997,** *15*, 1389-1395.

29. Hagestam, I. H.; Pinkerton, T. C. *Anal. Chem.* **1985,** *57*, 1757-1763.

30. Stout, R. W.; DeStefano, J. J. *J. Chromatogr.* **1983,** *261*, 189-212.

31. Wolters, D. A.; Washburn, M. P.; Yates, J. R., III *Anal. Chem.* **2001,** *73*, 5683-5690.

32. Gygi, S. P.; Rist, B.; Griffin, T. J.; Eng, J.; Aebersold, R. *J. Proteome Res.* **2002,** *1*, 47-54.

33. Wang, Q.; Svec, F.; Frechet, J. M. J. *Anal. Chem.* **1993**, 65, (17), 2243-2248.

34. Svec, F.; Frechet, J. M. J. *Anal. Chem.* **1992**, 64, (7), 820-822.

35. Luo, Q.; Shen, Y.; Hixson, K. K.; Zhao, R.; Yang, F.; Moore, R. J.; Mottaz, H. M.; Smith, R. D. *Anal. Chem.* **2005**, 77, (15), 5028-5035.

Chapter 4

Method Development and Applications of Extended pH-Range Chromatofocusing

4.1. Introduction

Two-dimensional polyacrylamide gel electrophoresis (2D-PAGE) is a very popular technique in protein analysis and is still one of the most widely used separation methods in proteome analysis.[1-5] However, this technique has some weaknesses. It is very labor-intensive and time-consuming. Its run-to-run reproducibility is generally not very good, although new differential gel electrophoresis can overcome this problem.[6, 7] Some proteins may be precipitated during electrophoresis resulting in low recovery. In addition, the gel technique is limited in dynamic range of detection. Thus the low abundance proteins are generally not observed.

About 10 years after the innovation of isoelectric focusing by electrophoresis (IEF) Sluyterman figured out a novel way to implement isoelectric focusing with a pH gradient on an LC column – it was called chromatofocusing.[8] Proteomic analysis methods that employ high performance liquid chromatography (HPLC) to achieve isoelectric focusing separations have been considered as an alternative to gel-based techniques, because of their improved run-to-run reproducibility and sensitivity.[9-11] Another advantage of chromatofocusing over gel-based techniques is that much larger sample quantities can be loaded for a single run. In addition, all fractions can easily be concentrated and subjected to buffer-exchange by using centrifugal filter devices. The most popular chromatofocusing method is based on the Polybuffer and Polybuffer exchangers from Amersham Pharmacia (Piscataway, NJ, USA).[12-16] A disadvantage of Polybuffer for preparative application is the removal of polyampholytes from the eluted protein fractions.[17] Moreover, the commercial chromatofocusing method's coverage of 3-6 pH units is insufficient to cover all proteins' pIs in one proteome analysis run.

This work is focused on developing a chromatofocusing technique having good resolution and an extended pH range for separating complex protein mixtures, such as human serum samples. In this work, the buffer used consisted of small molecules only. A pH gradient was established by adsorbing the small organic components onto an ion exchange column through an external gradient. The pH curves were found to be predictable. The linear pH range of 3 to 12 was achieved, which is wider than the commercially available or other reported chromatofocusing methods. This range would cover the pIs of almost all proteins in one run.[12-19] Finally, the extended pH-range chromatofocusing technique was used for milk and serum proteome analysis.

4.2. Experimental

4.2.1. Chemicals and Reagents

All organic solvents used were HPLC grade and were purchased from Fisher Scientific (Ottawa, ON, Canada). All other chemicals, reagents, serum, and protein standards were purchased from Sigma-Aldrich (Oakville, ON, Canada).

4.2.2. Sample Treatment

The serum was acidified to pH 2.0 with 10% trifluoroacetic acid (TFA) and centrifuged at 15000 × g for 30 min. The supernatant was used for injection. Skim milk, purchased from a local grocery store, was diluted 1:50 with water. Some of the milk casein was precipitated by acidifying the milk to pH 4.6 with 2 M acetic acid followed by centrifugation at 15000 × g for 20 min. Water used in this work was purified by a Milli-Q UV plus Ultrapure system (Millipore, Mississauga, ON) and was 18.3 MΩ cm in specific resistance or better.

4.2.3. Reversed Phase HPLC-ESI/ MS (or MS/MS)

An Agilent 1100 HPLC system with binary pump (Palo Alto, CA) was used for gradient generation and solvent delivery. A monolithic (poly(styrenedivinylbenzene) copolymer) capillary column, 5 cm × 200 μm i.d., from LC packings (Amsterdam, The Netherlands) was used for reversed phase HPLC-ESI-MS (or MS/MS). Mobile phase A consisted of 0.05% TFA in distilled water, and B consisted of 0.04% TFA in acetonitrile (ACN). All of the solvent composition ratios mentioned in this work are V:V ratios.

The injection amount is 1 μg of tryptic peptide. The gradient was set as 0 - 1 min from 5% B to 10% B, 1 - 45 min from 10% B to 35% B, 45-59 min from 35% B to 90% B, 59 - 60 min from 90% B to 5% B. Ion spectra were acquired on an LCQ Advantage ion trap mass spectrometer (ThermoFinnigan, San Jose, CA). Through Xcalibur the LCQ was set to acquire a full-scan mass spectrum between 400 and 1400 *m/z* followed by 4 data-dependent product ion mass spectral scans. The collision energy was set at 35%. Other parameters for Xcalibur were set as a duration of 0.5 min with 1 min exclusion window.

4.2.4. Microbore LC-MALDI QqTOF MS (or MSMS) for peptides

The peptide mixtures were separated by reversed phase chromatography on a 15 cm × 1.0 mm i.d. Vydac silica based C_{18} column with particle size of 5 μm and pore size of 300 Å (Hesperia, CA, USA) at a flow rate of 50 μL/min. The injection amount is 10 μg of tryptic peptide. Mobile phase A consisted of 0.1% TFA in distilled water, and B consisted of 0.1% TFA in ACN. An Agilent 1100 series capillary HPLC system with binary pump was used for gradient generation and solvent delivery. The gradient was set as 0 - 15 min from 0% B to 10% B, 15 - 80 min from 10% B to 35% B, 80-100 min from 35% to 80%, 100 - 105 min from 80% B to 0% B. Between 10 min and 100 min, 1 min fractions were collected onto a 100-well matrix-assisted laser desorption/ionization (MALDI) plate from Applied Biosystems (Concord, ON, Canada) through our home-built heated droplet LC-MALDI interface.[20] After fractionation, the dried peptides in each well were redissolved and mixed with 0.8 μL of 100 mg/mL DHB matrix solution in 50% methanol/50% water. Subsequent MALDI MS and MALDI MS/MS data were acquired on an Applied Biosystems/MDS-Sciex QSTAR Pulsar QqTOF instrument equipped with an orthogonal MALDI source employing a 337 nm nitrogen laser (Concord, ON, Canada). The instrument was operated in positive ion mode. The MS spectrum for each spot on the MALDI plate was collected in an automated mode and collision-induced dissociation (CID) of peptides was achieved with argon as collision gas.

4.2.5. MALDI MS for Proteins

MALDI experiments were performed using a Bruker REFLEX III MALDI/TOF (time-of-flight) mass spectrometer (Framingham, MA, USA) on each individual chromatofocusing fraction. Desorbed protein ions were produced with a 337 nm pulsed nitrogen laser and detected in the linear mode of operation in TOF. A two-layer sample preparation method with HCCA as matrix was used. The first layer of the two-layer method was performed by applying 0.8 μL of 0.1M HCCA in 10% methanol/90% acetone to the MALDI target. For the second layer, 0.1 μL solution from each chromatofocusing fraction was mixed with 1 μL of solution that consisted of saturated HCCA solution in 50% formic acid/acetonitrile/water (1:1:1), and applied onto the first layer and allowed to dry.

4.2.6. Database Search

MASCOT (Matrix Science Ltd) was used as search engine to perform peptide sequencing via database searching against the NCBI human database. The MS data processing criteria are described as follows:

(a) One MS/MS spectrum might have been assigned to more than one peptide candidate in Mascot. However, firstly, all non first candidates were not considered as possible matched peptides;

(b) For the first peptide candidates, their scores must have ensured that the probability for the peptide sequence matching a random a sequence was < 10%.

(c) The peptides could be assigned to more than one protein. Their first assignment candidates were the proteins that uniquely had the peptides. If >= 2 proteins had the same peptide, the peptide's first assignment candidate was the protein with the most peptides assigned or the protein with the highest scores;

(d) All identified proteins must have been at least one qualified peptide's first assignment candidate.

4.3. Results and Discussion

4.3.1. Internal and External Chromatofocusing

Chromatofocusing is a method used to achieve isoelectric focusing of analytes with a pH gradient produced during an LC separation. In chromatofocusing, a suitable ion-exchange column is pre-equilibrated with a starting buffer A to a certain pH, and then eluted with another buffer B at a different pH in 100% B isocratic or an A-B linear gradient to form a pH gradient on the column. The analytes are then selectively eluted at pH values proximal to their pI values. One major advantage of chromatofocusing arises from the fact that the process focuses the proteins to their pIs and simultaneously concentrates them as they are eluted from the column.

Two modes have been used to form a pH gradient along the column: the internal mode which is more widely used and the external mode. Initially the column is pre-equilibrated with a start buffer (A), and then in internal mode the column is eluted in an isocratic condition with 100% elution buffer (B), or in external mode the column is eluted in an A-B buffer gradient. The column used usually has buffer capacity: either a weak anion- or cation-exchange column. The workable pH range of existing commercial chromatofocusing is normally 3-6 pH units, limited by the pH range in which the column can keep its buffer capacity. However, chromatofocusing with this relatively narrow pH range is not adequate to separate complex protein mixtures where pIs of proteins generally range from 3 to 12.

In our work, a novel method is developed to establish a pH gradient on the column. It involves adsorption of compounds with buffer capacity along a strong cation exchange column, instead of a weak cation exchange column, followed by elution of the analytes through an external A-B buffer gradient condition. The buffer capacity on the column is created by the adsorbed buffering compounds, instead of the stationary phase. We call this method a hybrid external chromatofocusing. The schematic of the analytical setup is shown in Figure 4.1.

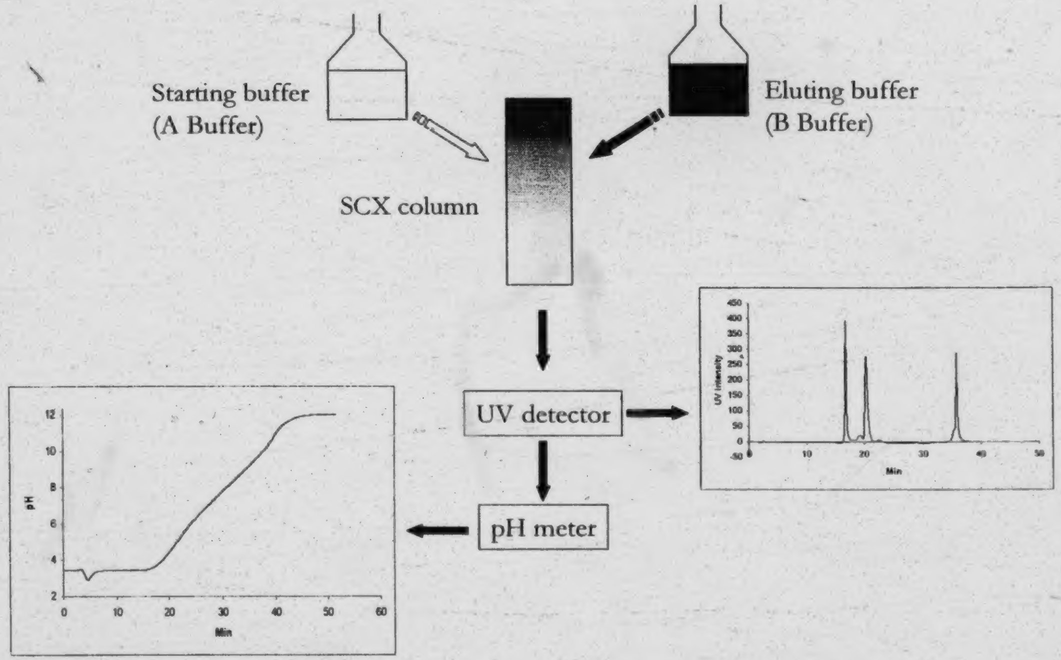

Starting buffer
(A Buffer)
Eluting buffer
(B Buffer)
SCX column
UV detector
pH meter
pH
12
10
8
6
4
2
0
10
20
30
40
50
60
Min
UV intensity
450
400
350
300
250
200
150
100
50
0
-50
0
10
20
30
40
50
Min

Figure 4.1. The schematic setup for the hybrid external chromatofocusing. The 2.1 mm i.d. strong cation exchange (the polystyrene-divinylbenzene based column with particle size of 10 μm and pore size of 1000 Å) and 1.0 mm i.d. strong cation exchange column (the polystyrene-divinylbenzene based column with particle size of 8 μm and pore size of 1000 Å) were purchased from Biochrom Labs (Terre Haute, IN, USA) and Vydac (Hesperia, CA, USA), respectively. The column was regenerated by flush the column with 20 column volumes of buffer A.

Figure 4.2 shows the titration curves of the 25 cm × 1.0 mm i.d. strong cation exchange column with different starting buffers: (I) was obtained by pre-equilibrating the column with 0.001 M HCl as starting buffer followed by titrating the column with 0.01 M

NaOH at a flow rate of 0.08 mL/min; (II) was obtained by pre-equilibrating the column with buffer (II) listed in Table 4.1 and titrating the column with 0.01 M NaOH at a flow rate of 0.08 mL/min. In Figure 4.2, curve I has a very sharp slope which indicates that the column's solid phase does not have buffer capacity between pH 3.0 and pH 12 and that the pH curve in this range is mainly determined by the pH's of the mobile phase and the gradient. The starting buffer for Figure 4.2 (I) does not contain any adsorbent components, while the starting buffer for Figure 4.2 (II) contains many components that can adsorb to the column and give the column buffer capacity.

Figure 4.2 (II) indicates that the strong cation exchange column can obtain different buffer capacity by adjusting the composition and pH of the starting buffer. In Figure 4.2, curve I has a greater slope in ΔpH/ΔV than curve II. Thus, curve I will have wider peak broadening in terms of pH, since the peak broadening in pH units is proportional to the square of ΔpH/ΔV.[8] The titration curves of the regular chromatofocusing methods are similar to curve I as the column used in the regular chromatofocusing methods has buffer capacity, but they may show different pH linearity in different pH ranges.

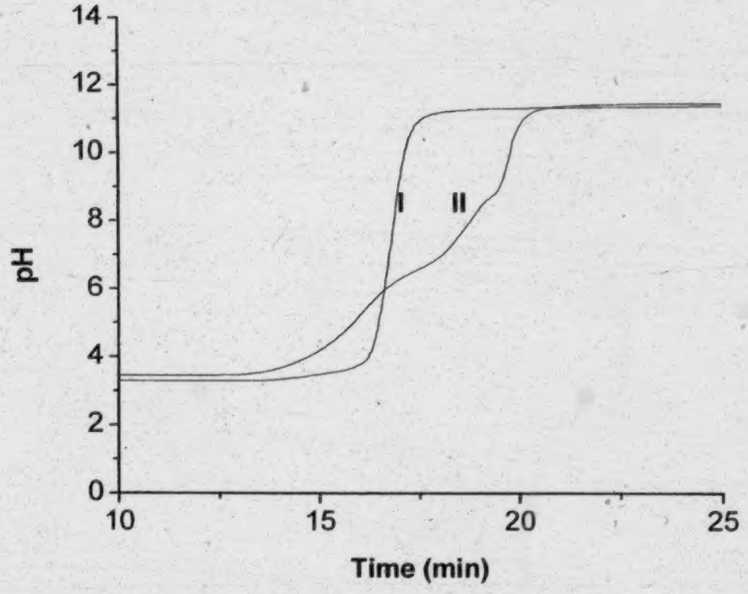


Figure 4.2. The titration curve of (I) 0.001 M HCl pre-equilibrated and (II) pre-equilibrated with buffer (B) listed in Table 4.1.

Table 4.1. Buffer composition (mM) used in this chapter. HCl and NaOH were used to adjust the pH of Buffer (I) or Buffer (II). The pH of buffer A with 2 M urea at around 3.0 and the pH of buffer B with 2 M urea at around 12 were used as the starting buffer and elution buffer, respectively.

	pKa	Buffer (I) (mM)	Buffer (II) (mM)
BICINE	8.35	1.25	2.5
Bis-Tris Propane	6.8; 9.0	1.25	2.5
CAPS	10.4	1.25	2.5
CHES	9.3	1.25	2.5
Citric Acid	3.13; 4.76; 6.40	2.5	1.25
Glycine	2.35; 9.78	1.25	2.5
Itaconic Acid	3.84; 5.55	2.5	1.25
Malic Acid	1.92; 6.23	2.5	1.25
Malonic Acid	2.85; 5.70	2.5	1.25
MES	6.21	1.25	2.5
MOPS	7.31	1.25	2.5
Oxalic Acid	1.25; 3.81	2.5	1.25
Phosphoric Acid	2.14; 6.86; 12.4	1.25	2.5
Succinic Acid	4.21; 5.64	2.5	1.25
Tartaric Acid	3.17; 4.91	2.5	1.25
TRICINE	8.26	1.25	2.5
Triethanol Amine	7.8	1.25	2.5
Triethylamine	10.75	1.25	2.5
Tris (hydroxymethyl) Amino-Methane	8.1	1.25	2.5

4.3.2. pH Linearity of the Hybrid External Chromatofocusing

The buffer capacity (B_s) of the column in the hybrid external chromatofocusing is a function of the total buffer capacity of the adsorbents (B_a) and the total buffer capacity of cation exchange functional groups on the column (B_c):

$$B_s = B_c + B_a \tag{4.1}$$

From Figure 4.2 (I) it can be seen that the total buffer capacity of the cation exchange functional group B_c can be ignored, and that the buffer capacity B_s is roughly equal to the total buffer capacity of the adsorbents on the column B_a, i.e., the additive buffer capacity of the adsorbents on the column:

$$B_a = \sum_{i=1}^{N} B_i \tag{4.2}$$

where B_i is the buffer capacity of the adsorbent i on the column and N is the total number of the adsorbents on the column.

In equation (4.2), B_i can be expressed as:

$$B_i = 2.3 \times C_{i,s} \times a_{i,s} \times (1 - a_{i,s}) \tag{4.3}$$

where $C_{i,s}$ is the concentration of the adsorbent *i* on the column and $a_{i,s}$ is the distribution function of the conjugate acid form of the adsorbent *i* on the column.

The cation exchange column is negatively charged between pH 3 and pH 12. Assuming only the positively charged adsorbents to be adsorbed onto the column, then the following equation can be used to describe the charged adsorbents between the stationary phase and mobile phase[8]:

$$C_{i,s} = C_{i,m} \times e^{\frac{F\varphi}{RT}Z} \tag{4.4}$$

where $C_{i,s}$ and $C_{i,m}$ are the concentrations of adsorbent *i* on the stationary phase and in the mobile phase respectively, φ is the Donnan potential, and *F*, *R*, and *T* have their conventional meanings. Inserting equation 4.2-4.4 into 4.1, gives:

$$B_s = \sum_{i=1}^{N} 2.3 \times C_{i,m} \times e^{\frac{F\varphi}{RT}Z} \times a_{i,s} \times (1 - a_{i,s}) \tag{4.5}$$

When the fraction *j* mobile phase flows from position *k* to position *k+1* on the column, after mixing with the stationary phase its pH change $dpH_{m,j,k+1}$ is:

$$dpH_{m,j,k+1} = pH_{m,j,k+1} - pH_{m,j,k} \tag{4.6}$$

where $pH_{m,j,k+1}$ and $pH_{m,j,k}$ are the pH of the fraction *j* mobile phase on position *k+1* and *k* along the column, respectively.

$$dpH_{m,j,k+1} = \frac{dX_{b,m}}{B_m} \tag{4.7}$$

where $dX_{b,m}$ is the amount of base removed from the fraction *j* mobile phase and B_m is the buffer capacity of the mobile phase. Inserting equation 4.7 into 4.6:

$$pH_{m,j,k+1} - pH_{m,j,k} = \frac{dX_{b,m}}{B_m} \tag{4.8}$$

After mixing with the fraction *j* mobile phase, the pH change of the stationary phase at position *k+1* $dpH_{s,j,k+1}$ can be expressed as:

$$dpH_{s,j,k+1} = pH_{s,j,k+1} - pH_{s,j-1,k+1} \tag{4.9}$$

where $pH_{s,j,k+1}$ *and* $pH_{s,j-1,k+1}$ are the pH when the fraction *j* and *j-1* mobile phase pass over the stationary phase at position *k+1*.

$$dpH_{s,j,k+1} = \frac{dX_{b,s}}{B_s} \tag{4.10}$$

where $dX_{b,s}$ is the amount of base added into the stationary phase at position *k+1*. Inserting equation 4.10 into 4.9:

$$pH_{s,j,k+1} - pH_{s,j-1,k+1} = \frac{dX_{b,s}}{B_s} \tag{4.11}$$

Assuming that after acid-base mixing the pH of the acid and the pH of the base are equal, i.e.,:

$$pH_{s,j,k+1} = pH_{m,j,k+1} \tag{4.12}$$

$$pH_{s,j-1,k+1} = pH_{m,j-1,k+1} \tag{4.13}$$

After the fraction *j* mobile phase mixes with the stationary phase at position *k+1*, the amount of the base added into the stationary phase, $dX_{b,s}$, is equal to the amount of base removed from the fraction *j* mobile phase $dX_{b,m}$:

$$dX_{b,s} = -dX_{b,m} \tag{4.14}$$

Inserting equation 4.12 – 4.14 into 4.11, gives:

$$pH_{m,j,k+1} - pH_{m,j-1,k+1} = \frac{-dX_{b,m}}{B_s} \tag{4.15}$$

Defining:

$$R_c = \frac{B_s}{B_m} \tag{4.16}$$

and solving equation 4.14 and 4.15, gives:

$$pH_{m,j,k+1} = \frac{pH_{m,j,k} + R_c \times pH_{m,j-1,k+1}}{1 + R_c} \tag{4.17}$$

B_m in equation 4.16 can be calculated as:

$$B_m = 2.3 \times C_{i,m} \times a_{i,m} \times (1 - a_{i,m}) \tag{4.18}$$

, where $a_{i,m}$ is the distribution function of the conjugate acid form of the adsorbent *i* on the column. The difficulty in using equation 4.5 to calculate B_s is that the Donnan potential cannot be easily estimated on the column. To make things easier in the simulation, equation 4.5 was simplified as:

$$B_s = \sum_{i=1}^{N} \alpha_i \times C_{i,m} \tag{4.19}$$

where α_i is the constant related to the adsorbent i on the column.

Figure 4.3 (I) and (III) shows the pH simulation results based on equation 4.19 when $\alpha_i = 0.3$ was assumed. (II) and (IV) are the experimental pH curves on a 25 cm × 1.0 mm i.d. column when buffers (I) and (II) listed in Table 4.1 were applied, respectively. The simulation results agree well with the experimental results. This means that the model can be used to predict the pH linearity when different buffer compositions are to be used. Buffer (II) listed in Table 4.1 provides better linearity and its pH range is wider than that of any existing chromatofocusing systems. The chromatofocusing with this wide pH range is adequate to separate complex protein mixtures where pI values of proteins generally range from 3 to 12.

Buffer (II) in Table 4.1 was selected as the running buffer in the following studies in this chapter unless individually described.

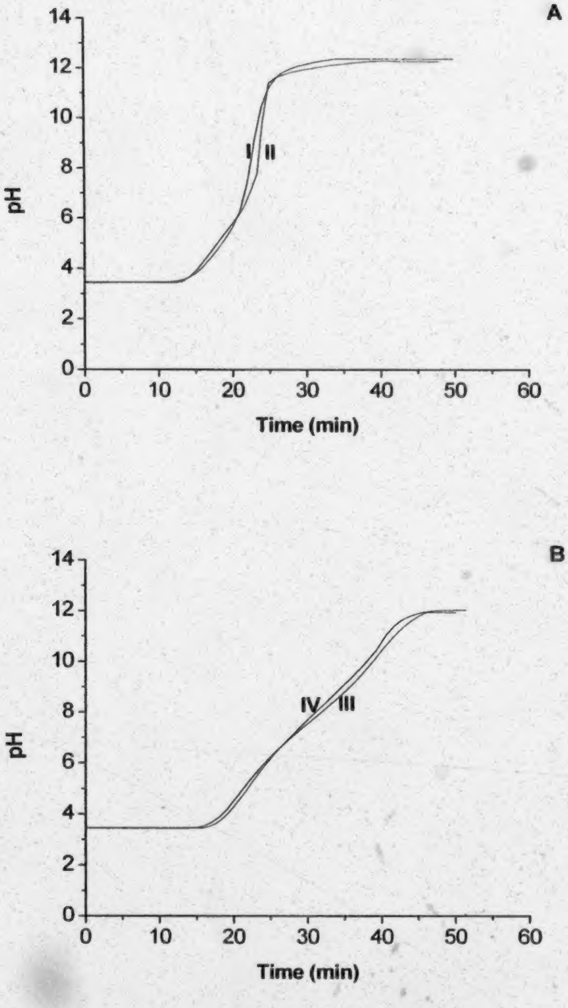


Figure 4.3. (I) and (III) are the pH simulation results based on equation 4.17 when α_i= 0.3 was assumed in equation 4.18 when buffer (I) and (II) in Table 4.1 were applied respectively. (II) and (IV) are the experimental pH curves on 25 cm × 1.0 mm i.d. column when buffer (I) and (II) in Table 4.1 were applied respectively. The flow rate was 0.08 mL/min.

The pH linearity is very important for the resolution of chromatofocusing. Figure 4.4 (A) and (B) were obtained when buffer (I) and (II) in Table 4.1 were used as the running buffer respectively. The other HPLC conditions were the same: the 15 cm × 2.1 mm i.d. strong cation exchange column was used; A 30 min A-B linear gradient was used; the flow rate was 0.2 mL/min; and the detection wavelength was 280 nm with reference wavelength at 360 nm. 30 μg trypsinogen, 25 μg α-chymotrypsin and 10 μg lysozyme were injected in each run.

It can be seen that the dpH/dt vs time curve, i.e., the slopes of the pH curves in Figure 4.4 (A) and (B) are different: the slope of the pH curve in Figure 4.4 (B) is flat while it goes up and down in Figure 4.4 (A) and the slope of the pH curve in Figure 4.4 (A) is much greater than that in Figure 4.4 (B). The difference in pH curve results in a performance difference of the two running buffers in separation: the peaks came out earlier in Figure 4.4 (A) and the separation factors between trypsinogen and α-chymotrypsin are 1.1 and 1.6 in Figures 4.4 (A) and (B), respectively. Obviously buffer (II) in Table 4.1 led to better resolution than buffer (I).

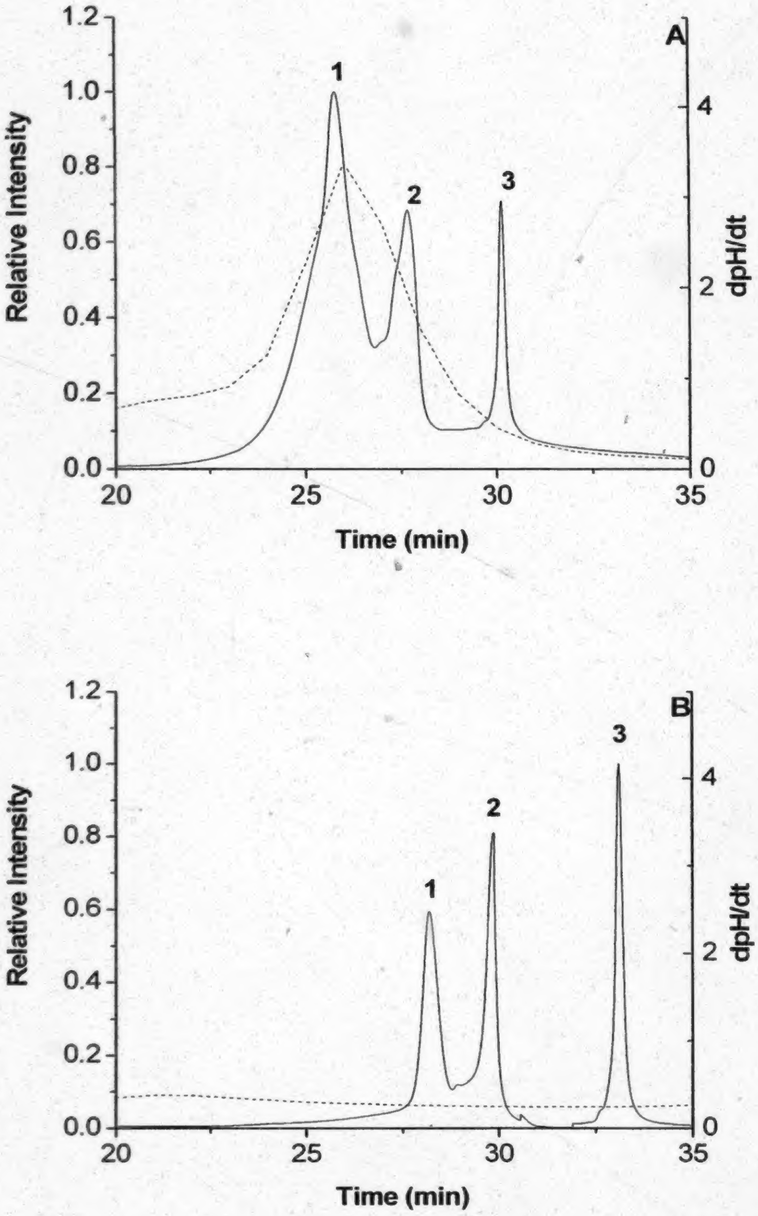


Figure 4.4. Chromatograms obtained when buffer (I) and (II) in Table 4.1 were used as the running buffers. The other HPLC conditions were the same: a 15 cm × 2.1 mm i.d. strong cation exchange column was used; the flow rate was 0.2 mL/min; the detection wavelength was 280 nm with reference wavelength at 360 nm. 25 μg α-chymotrypsin (1), 30 μg trypsinogen (2) and 10 μg lysozyme (3) were injected in each run. A 30 min A-B linear gradient was used.

4.3.3. Mechanism of Chromatofocusing

Figure 4.5 shows the computational results of pH change based on equation 4.19 when buffer aliquots are run through the column. The start buffer pH and the elution buffer pH were supposed to be 3.3 and 11.7, respectively. In Figure 4.5 the dashed lines from bottom to top present the pH trends of successive aliquots of the elution buffer passing over the axis of the column. Because the column was pre-saturated with the starting buffer and eluted in an A-B linear gradient, the first aliquot of the elution buffer was the least changed aliquot. The column was a strong cation exchange column and was negatively charged in pH range from starting buffer pH 3.3 to elution buffer pH 11.7. The protein pI was supposed to be 8.0. The pH of the starting buffer was 3.3, so the protein was positively charged and adsorbed on the column at the beginning of the separation. When the first aliquot of the elution buffer with pH > 8.0 flowed through the column, the protein was negatively charged since the pH was greater than the protein's pI, and was dissolved into this aliquot of elution buffer until this aliquot of elution buffer was neutralized along the column and its pH became lower than the protein's pI. At this point the protein was positively charged again and adsorbed on the column. When the following aliquot of the elution buffer passed over the protein, its pH was higher than the protein's pI and took the protein moving forward until this aliquot of the elution buffer was neutralized along the column and the pH became lower than the protein's pI again. Other following aliquots of the elution buffer took the protein moving forward successively in this way until the protein was eluted, i.e., the pH of the elution buffer at the end of the column was close to 8.

In the ascending pH gradient, as shown in Figure 4.5, the protein was constantly and cyclically changing its net surface charge from positive charge, zero charge and to negative charge as the pH gradient developed and the protein traveled through different pH zones on the column. Molecules at the rear of the protein band always migrated faster than those at the front because the molecules at the rear were more negatively charged and more repulsed from the column. Gradually in this way a narrower protein band was formed along the column. The saw-tooth wave in Figure 4.5 presents the focusing path. After every stepwise pH change, the band of the protein was focused to a small extent.

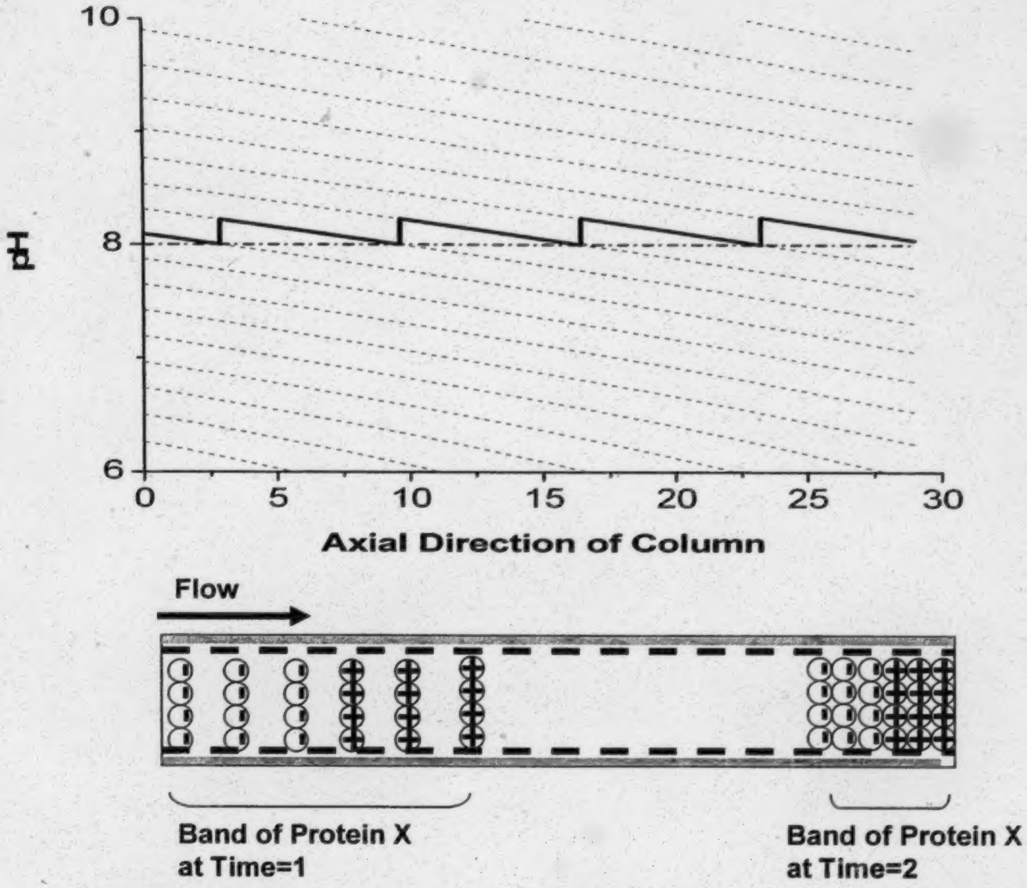


Figure 4.5. Computational simulation of the pH changes of the elution buffer passing over the column (the dashed lines) and the focusing path of the protein X with pI 8.0 (the saw-tooth wave). The pH of the start buffer is 3.3, and the pH of the elution buffer is 11.7. The focusing effect led to narrower band width of protein X.

The experimental data presented in Figure 4.6 show the effect of column length on chromatofocusing separation. In Figure 4.6, (A) and (B) resulted from the same HPLC conditions of separation except the column length was different. The column length in (A) was longer than that in (B), so (A) should have a longer focusing path and better focusing effect based on the simulation results shown in Figure 4.5. When the column length decreased, the possible equilibrations of an aliquot of elution buffer and the stationary phase decreased. Then the chance of the dissolved protein in an aliquot of elution buffer to bind back with the stationary phase was reduced and a portion of the protein came out before its pI. So the recovery from the shorter column was worse.

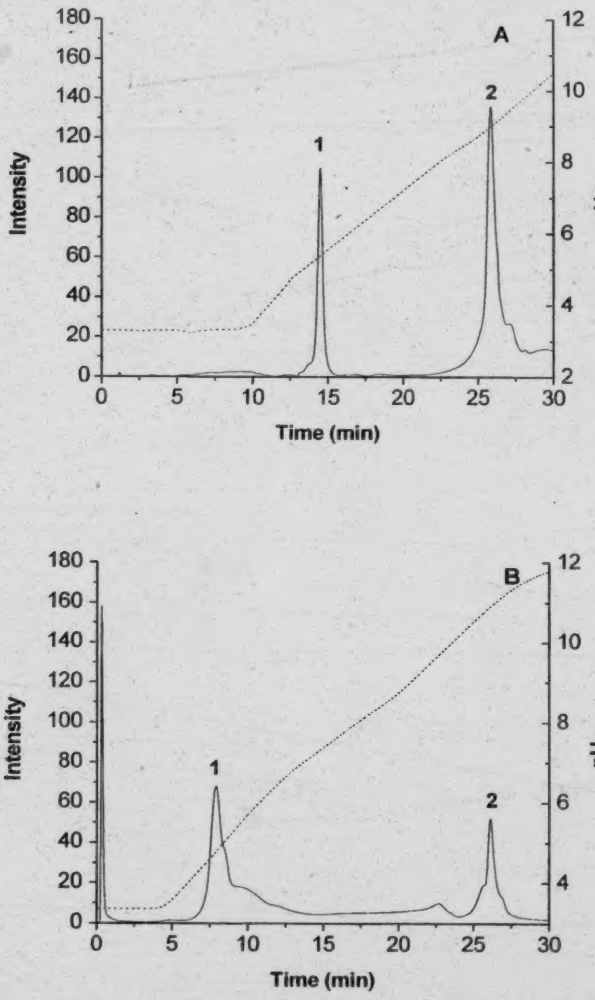


Figure 4.6. Chromatograms (A) and (B) obtained when buffer (II) in Table 4.1 was used as the running buffer on a 15 cm × 2.1 mm i.d. strong cation exchange column and a 2.5 cm × 2.1 mm i.d. strong cation exchange column, respectively. The other HPLC conditions were the same: the flow rate was 0.2 mL/min; the detection wavelength was 280 nm with reference wavelength at 360 nm. 25 μg β-lactoglobulin A (1) and 30 μg trypsinogen (2) were injected in each run. A 20 min A-B linear gradient was used.

4.3.4. Analysis of Standard Proteins with Extended pH-range Chromatofocusing

Figure 4.7 illustrates the separation of a mixture of 6 standard proteins using the extended pH-range chromatofocusing method. These standard proteins with pI values ranging from 4.0 to 11.0 can not be separated by traditional chromatofocusing methods that have a narrow range of 3 to 6 pH units. Proteins with pI outside this narrow range will not be retained or eluted from the column. In the latter case a salt gradient has to be applied after chromatofocusing.[21]

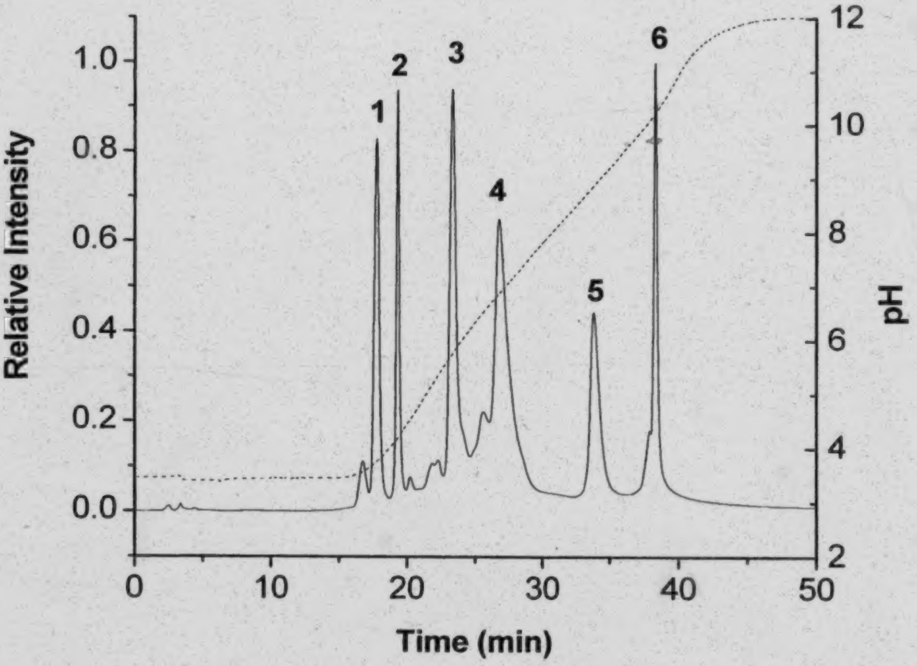


Figure 4.7. Chromatogram of 6 standard proteins separated with the extended pH-range chromatofocusing method. Buffer (II) in Table 4.1 was used as the running buffer on a 25 cm × 1.0 mm i.d. strong cation exchange column. The flow rate was 0.08 mL/min; the detection wavelength was 280 nm with reference wavelength at 305 nm. 60 μg ovalbumin (1), 20 μg β-lactoglobulin A (2), 30 μg α-chymotrypsin (3), 25 μg trypsinogen (4), 60 μg cytochrome C (5), and 10 μg lysozyme (6) were injected in each run. A 30 min A-B linear gradient was used.

Based on the discussion on focusing mechanism simulation, the elution of the proteins would follow the order of their pIs from low to high. However, a protein's observed pI, i.e., the pH at which the protein is eluted, is somewhat different from its theoretical pI. Besides the theoretical pIs, the ionic strength and composition of the running buffer, the stationary phase, the flow rate and the temperature can affect a protein's observed pI because these factors can change a protein's folding and its interaction with the column. Table 4.2 is the summary of the 6 proteins' theoretical pIs and observed pIs in Figure 4.7.

Table 4.2. The 6 proteins' theoretical pIs and observed pIs in Figure 4.7.

	Theroretical pI	Observed pI
Ovalbumin	4.7	4.0
β-lactoglobulin A	5.1	4.3
α-chymotrypsin	8.5	5.7
Trypsinogen	9.3	6.9
Cytochrome C	10.6	8.9
Lysozyme	11.0	10.4

The difference between the observed pI and the theoretical pI, ΔpI, can be roughly described by equation 4.19:[8]

$$\Delta pI = \frac{1}{4.6}\varphi + \frac{\dfrac{k}{dZ/dpH}}{\varphi} \quad (4.19)$$

where Z is the charge of the protein, φ is the Donnan potential and k is related the buffer capacity of the column and the running buffer. dZ/dpH in equation 4.19 was different for different proteins. In Figure 4.17 it can be seen that α-chymotrypsin and trypsinogen had the widest peak width, meaning that these two proteins had the smallest dZ/dpH, since the peak width is inversely proportional to the square of dZ/dpH.[8] In equation 4.19 it can be seen that a smaller dZ/dpH can lead to a greater ΔpI. This reduction agrees with the fact that in Table 4.2 α-chymotrypsin and trypsinogen had the highest ΔpIs. The other two parameters in equation 4.19, K and φ, also changed when pH changed on the column. Since all parameters in equation 4.19 could be different for different proteins, it is not surprising to see different proteins having different ΔpIs. Therefore it is important to keep using the same column and the same buffer at the same temperature when pIs of the proteins are to be evaluated with the chromatofocusing method. Evaluating protein pIs with chromatofocusing has been practiced for many years. Sometimes if the proteins' pIs are too different to be covered by one chromatofocusing method, different buffers and columns have to be applied.[22-24] However, the extended pH-range chromatofocusing described here offers a better choice since it can cover almost all proteins' pIs; thus making a buffer or column switch unnecessary.

Chromatofocusing is usually performed with very low ionic strength in the running buffer. But it is still necessary to adjust the pH with charged counter ions such as Cl^- or K^+. These ions bring additional ionic strength that causes proteins to elute prematurely, i.e., below their theoretical pI. From Table 4.2 it can be seen that the observed pIs of all 6 standards are less than their theoretical pIs. The inconsistency between the observed pIs and the theoretical pIs can also be related to a geometrical factor between the proteins and the column. The proteins have three-dimensional structures. Only a portion of a protein surface is bound to the surface of the stationary phase. Because the stationary phase is negatively charged, the portion of the protein's surface binding to the stationary phase is the part most liable to be positively charged. If a protein's charges are localized evenly on the protein's surface, the protein's elution behavior would likely follow the order of its true pI. However, if the protein's charges tend to be distributed unevenly, the surface part that is most prone to be positively charged should be positively charged before the whole protein appears to be positively charged. In this case the pI of the contact surface (here called contact pI) should be less than the whole protein's surface pI. (it should be mentioned that the protein's surface pI may also be different from the protein's theoretical pI since the surface pI depends on the protein's folding and unfolding.) This phenomenon is related to the protein orientation preference and its practical significance is that a protein's contact pI may be significantly different from its true surface pI, not to mention the theoretical pI. Elution behaviors of a protein in different HPLC conditions will be somewhat different, so the observed pI, i.e. the contact pI, will deviate accordingly. The developed method has extended the pH range to 3-12, which can cover almost all proteins, so it is a better choice for experimentally evaluating some important protein's pI than the commercial chromatofocusing methods with a pH range of 3-6 units. However, the pI comparison is valuable only when the same HPLC conditions are used.

Tables 4.3 and 4.4 show the effects of the flow rate and the gradient on the observed pI, respectively. Using a slower flow rate or gradient, the observed pI is a little bit more close to the theoretical pI. A possible explanation for this is that the protein's folding and unfolding becomes easier and more comprehensive at a slower flow rate or gradient, resulting in greater dZ/dpH in equation 4.19.

Table 4.3. Influence of the flow rate f on the observed pI. The A-B linear gradient was 30 min. The column was 25 cm × 1.0 mm i.d.

	β-lactoglobulin A	α-chymotrypsin	Lysozyme
Observed pI (f=0.12ml/min)	4.20	5.60	10.21
Observed pI (f=0.10ml/min)	4.30	5.68	10.33
Observed pI (f=0.08ml/min)	4.35	5.69	10.39

Table 4.4. Influence of the gradient t_G on the observed pI. The flow rate was 0.1 mL/min. The column was 25 cm × 1.0 mm i.d.

	β-lactoglobulin A	α-chymotrypsin	Lysozyme
Observed pI (t_G=24min)	4.11	5.66	10.21
Observed pI (t_G=30min)	4.30	5.68	10.33
Observed pI (t_G=45min)	4.41	5.86	10.46

To compare this new chromatofocusing technique with ion exchange methods, Figure 4.8 illustrates that separation efficiency of the method is higher than that of ion-exchange chromatography. The half peak widths of the three standard proteins, β-lactoglobulin A, trypsinogen, and lysozyme in Figure 4.8 are: 0.75, 0.84, 1.13 min, respectively, in (A) and 0.24, 0.50 and 0.25 min, respectively, in (B). The separation factors between β-lactoglobulin A and trypsinogen and between trypsinogen and lysozyme are: 1.4 and 1.5, respectively, in (A) and 2.3 and 10.33, respectively, in (B).

One of the noticeable characteristics of chromatofocusing is that chromatofocusing is conducted at low ionic strength. This prevents salt-elution effects from skewing the selectivity that is observed in the salt-gradient cation exchange chromatography. In contrast to the salt-gradient cation exchange, where peak width has an unfortunate tendency to broaden with elution time, as shown in Figure 4.8 (A), chromatofocusing peaks tend to remain very sharp due to the characteristic focusing effect. The column regeneration time for the chromatofocusing is found to be 3 times shorter than the strong cation exchange method and this can also be attributed to the low ionic strength in the running buffer. In addition, the use of relatively low salt concentration is particularly important for downstream MS analysis or for combining with other HPLC methods. The protein recovery shown in Figure 4.8 indicates that the chromatofocusing method and the strong cation exchange method has similar performance.

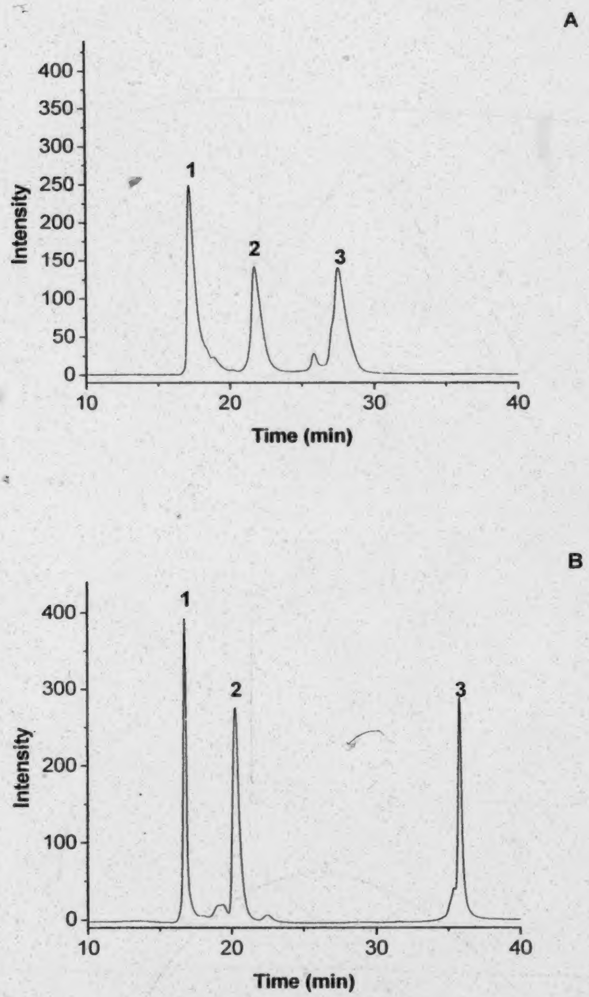


Figure 4.8. Chromatogram of 3 standard proteins separated with (A) strong cation exchange and (B) the extended pH-range chromatofocusing method on the same column: 25 cm × 1.0 mm i.d. strong cation exchange column and at the same flow rate: 0.10 mL/min. Buffer (II) in Table 4.1 was used for the chromatofocusing. 10 mM phosphate buffer at pH 3.0 was used as the running buffer for the cation exchange separation. 0.5 M KCl was added into the buffer B for the cation exchange. The other HPLC conditions were the same: detection wavelength was 280 nm with reference wavelength at 360 nm; 40 μg β-lactoglobulin A (1), 48 μg trypsinogen (2), and 15 μg lysozyme (3) were injected; A 30 min A-B linear gradient was used.

4.3.5. Application of Chromatofocusing on Milk Separation

To demonstrate the applicability of the extended pH-range chromatofocusing method to real world samples, it was used to separate the proteins in milk. Figure 4.9 shows the chromatogram of 50 μL pretreated milk (600 μg protein by protein assay) separated on a 15 cm × 2.1 mm i.d. strong cation exchange column. The buffer (II) listed in Table 4.2 was used as the running buffer for the chromatofocusing separation. As it is shown in Figure 4.9, 10 fractions were collected for MALDI-TOF MS described in Section 4.2.5. Figure 4.10 show the MS spectra of the 10 fractions plus the MS spectrum of the milk sample without the chromatofocusing separation. Masses of the proteins detected in the MS spectra (S/N > 5) in Figure 4.10 are listed in Table 4.5. Because of the ion suppression in the MALDI MS analysis of non-separated milk, only 14 most abundance proteins can be detected in the milk. This number is only one-tenth of the number of the proteins detected from the 10 chromatofocusing separated fractions. Figure 4.10 also show that the MALDI MS spectral patterns of individual fractions are different. Most proteins were detected in one fraction and only 11.5% of the detected proteins were found in two or more fractions. These results indicate the chromatofocusing separation method provides high efficiency for separating these milk proteins.

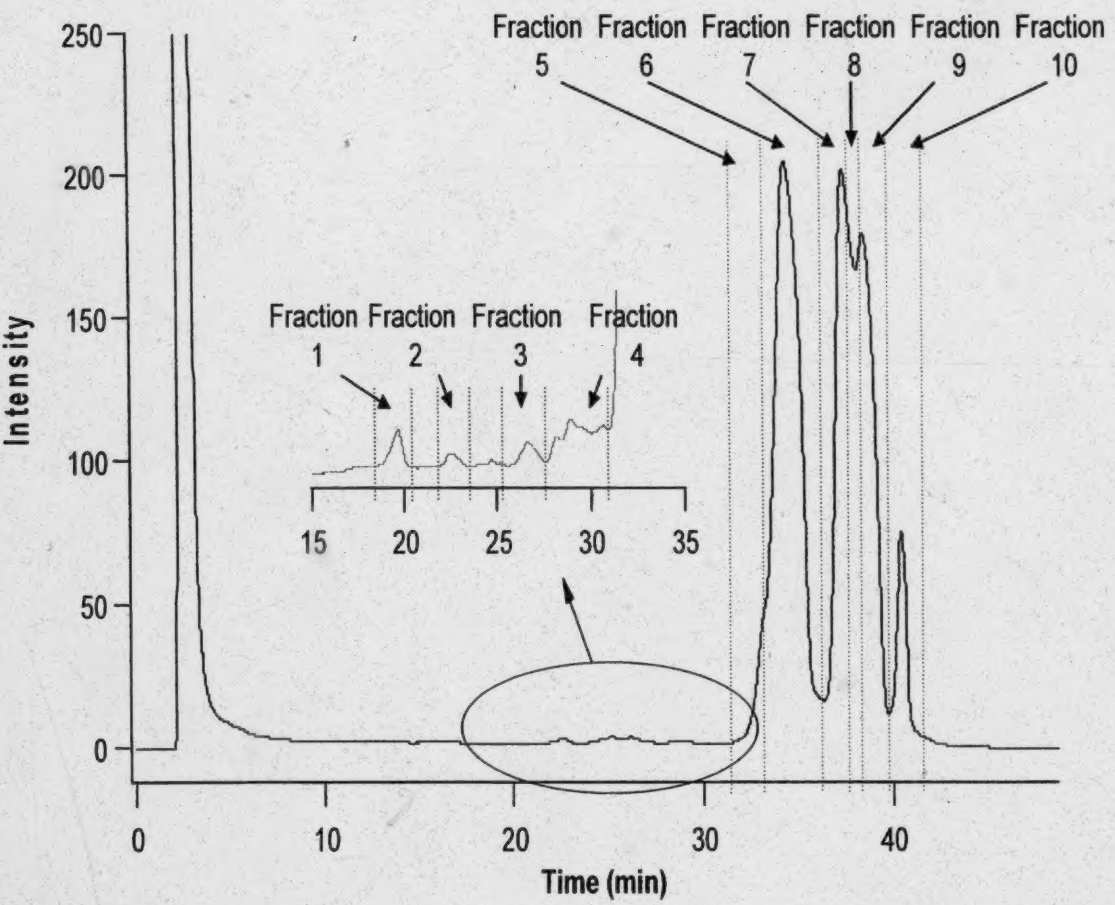


Figure 4.9. The chromatogram of 50 μL pretreated milk separated with the extended pH-range chromatofocusing. The column was 15 cm × 2.1 mm i.d. strong cation exchange column. The flow rate was 0.2 mL/min and a 50 min A-B linear gradient was applied. The detection wavelength was the same: 280 nm with reference wavelength at 360 nm. The buffer (II) listed in Table 4.2 was used as the running buffer for the chromatofocusing separation. 10 fractions were collected for MALDI/MS.

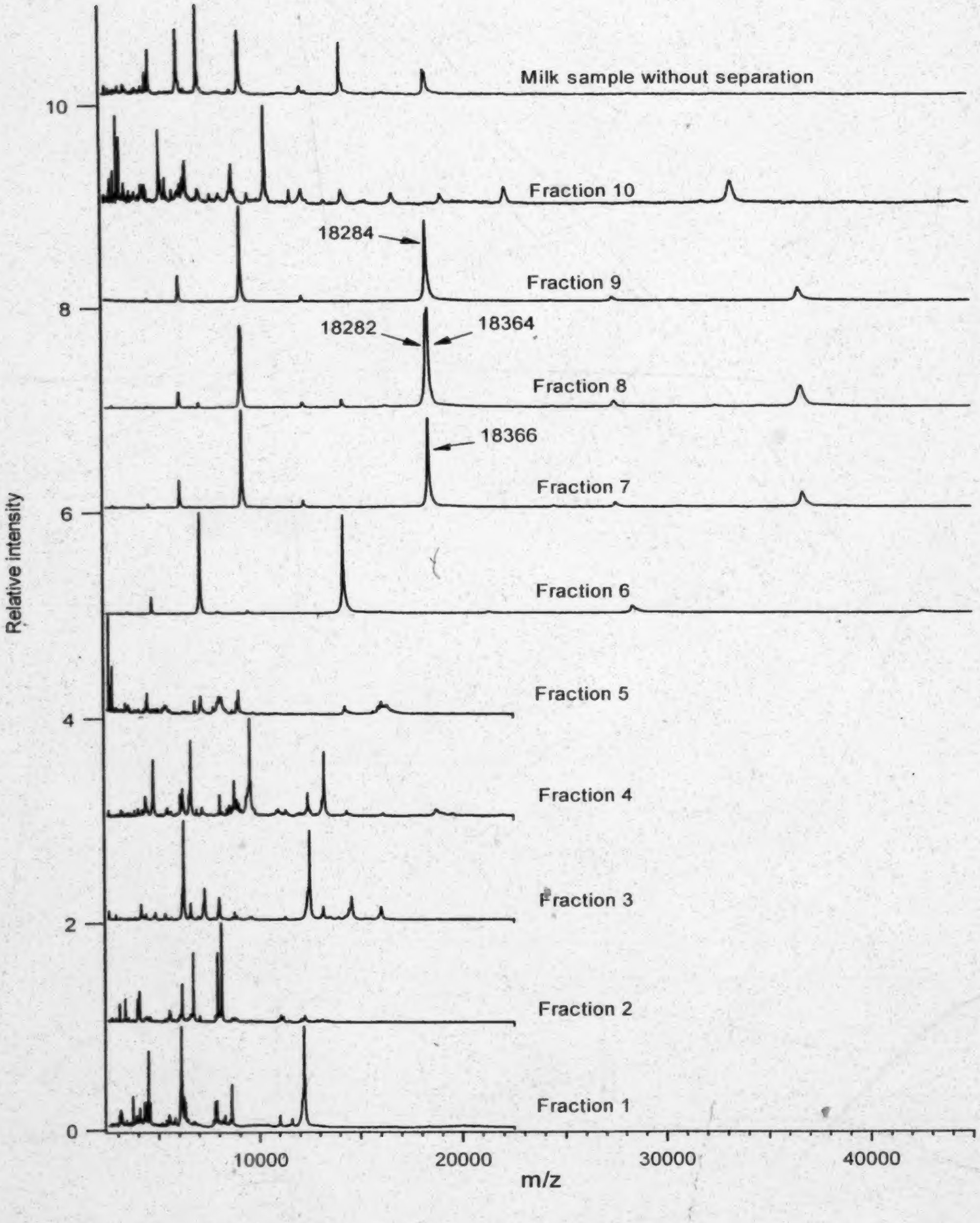


Figure 4.10. MALDI MS spectra of the 10 fractions plus the MALDI MS spectrum of the original milk sample with no fractionation.

Table 4.5. Masses of proteins identified in the MALDI MS spectra (S/N > 5) shown in Figure 4.10.

	Proteins/Peptides (MH+, Da)
Milk	2335.1, 2461.5, 2618.3, 2766.0, 2830.0, 3551.4, 3986.2, 4067.5, 4371.8, 6383.25*, 12209, 14182, 18284, 18364
Milk Separated with Chromatofocusing	
Fraction #01	2181.4, 2241.4, 2292.4, 3052.8, 3195.1, 3396.4, 3509.9, 3624.2, 3723.1, 4374.0, 4501.8, 4573.6, 5074.1, 5168.3, 5360.7, 5547.3, 6222.4, 6301.9, 6695.3, 6776.3, 7688.4, 7774.9, 7863.1, 8258.7, 8589.8, 10996, 11590, 12179
Fraction #02	2692.7, 2891.3, 4497.5, 4562.5, 5316.6, 6142.7, 6694.1, 7028.4, 7783.0, 7903.8, 8003.1, 8091.1, 8618.4, 8718.5, 11049, 11150, 12228, 13114
Fraction #03	2338.3, 2429.4, 2547.8, 2907.5, 4216.0, 4294.6, 8759.3, 8855.8, 9518.7, 9606.6, 11263, 12487, 13177, 14557, 16002
Fraction #04	2491.8, 2565.1, 3274.2, 3853.7, 4218.5, 6089.5, 6907.7, 7169.8, 8052.9, 8188.9, 8407.9, 8527.3, 8683.3, 8760.5, 8897.2, 8987.4, 9528.6, 10928, 11306, 12408, 13220, 14340, 16111, 18710
Fraction #05	2180.0, 2276.4, 2321.1, 2390.8, 2586.7, 2700.1, 2767.8, 3567.97*, 3668.2, 4069.1, 4722.0, 4859.4, 4971.8, 5030.0, 6827.6, 7744.7, 8682.0, 8898.4, 8988.0, 14270, 16049
Fraction #06	3543.8, 4523.1, 6220.0, 8014.0, 9028.5, 9463.4, 14186, 21336
Fraction #07	12253, 18366, 27577, 45980
Fraction #08	2827.4, 12229, 14182, 18283, 18364, 25487, 27518, 30612, 32542, 45882, 54999
Fraction #09	4584.3, 12198, 18284, 27461
Fraction #10	2106.8, 2460.2 2808.0, 2957.2, 3120.7, 3249.0, 3985.5, 4309.33*, 4497.8, 4574.8, 5316.2, 5491.8, 6367.9, 6469.3, 7709.2, 8158.0, 8763.2, 10416, 11641, 12209, 12214, 14223, 19084, 66534

4.3.6. Application of Chromatofocusing on Human Serum Separation

Serum is attracting an increasing interest in proteomics, which is currently striving to determine biomarkers through which the onset of a diseased state can be effectively distinguished. Serum has high protein content because serum constantly suffuses tissues from which many secreted proteins influx into the serum. However, high abundance proteins like albumin and transferrin dominate the protein content of serum, and the protein concentration dynamic range can span as high as 14 orders of magnitude.[25] Though state-of-art mass spectrometry (MS) provides an invaluable detection technique in efficiency and accuracy, characterization of proteins in such a huge dynamic range depends greatly on the ability to separate the proteins prior to MS detection. Many attempts of combining gel and chromatographic techniques have been made on serum separation.

Figure 4.11 shows the chromatogram of 50 µL pre-treated human serum separated on a 15 cm × 2.1 mm strong cation exchange column. The buffer (II) listed in Table 4.1 was used as the running buffer of the chromatofocusing. It can be seen that most of the proteins elute out between 28 min and 39 min. This elution window is mainly from the most dominant protein, albumin. Thus, the chromatofocusing method can be used for depleting high abundance proteins, like albumin, in human serum, if only the proteins eluted outside this window are collected.

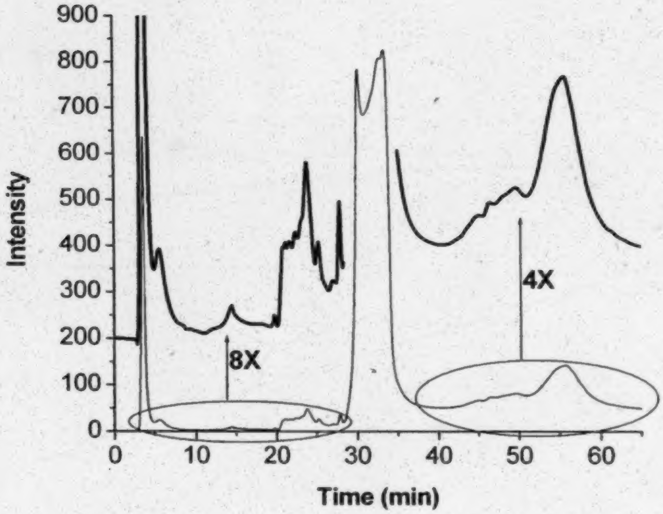


Figure 4.11. The chromatogram of 50 μL pretreated human serum separated with the pH-range extended chromatofocusing. The column was a 15 cm × 2.1 mm i.d. strong cation exchange column. The flow rate was 0.2 mL/min and a 60 min A-B linear gradient was applied. The detection wavelength was the same: 280 nm with reference wavelength at 360 nm. The buffer (II) listed in Table 4.2 was used as the running buffer for the chromatofocusing separation.

To evaluate the reproducibility of albumin depletion, 4 repeat runs were performed at an interval of 7 days. The HPLC conditions were the same as those in Figure 4.11. In each run the fraction between 28 min and 39 min was excluded and the other fractions were mixed.

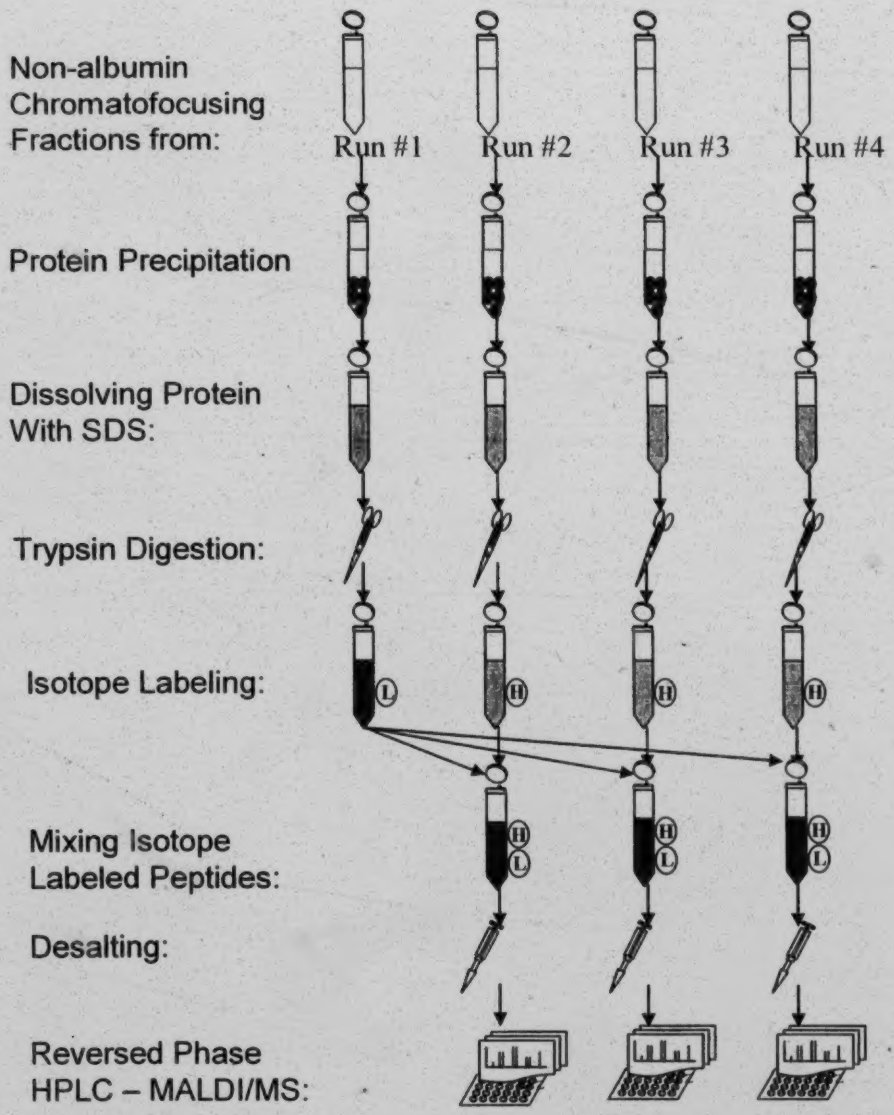

Non-albumin Chromatofocusing Fractions from:
Run #1
Run #2
Run #3
Run #4
Protein Precipitation
Dissolving Protein With SDS:
Trypsin Digestion:
Isotope Labeling:
L
H
H
H
Mixing Isotope Labeled Peptides:
H
L
H
L
H
L
Desalting:
Reversed Phase HPLC – MALDI/MS:

Figure 4.12. Schematic setup for a reproducibility investigation regarding albumin depletion with extended pH-range chromatofocusing.

Figure 4.12 shows the protocols for a reproducibility study of albumin depletion with the extended pH-range chromatofocusing. The four protein collections are labeled (I), (II), (III) and (IV), respectively, in the order of experimental date. The proteins in the

4 collections were precipitated by adding cold (-80 °C) acetone. The final ratio of acetone in the solution was 80%. Then the 4 solutions were kept at -20 °C overnight. They were centrifuged at 3000 x g in a Beckman J2-21 centrifuge (Mississauga, ON, Canada) at 4 °C. The supernatants were removed and the protein precipitates were dissolved in 0.01% sodium dodecylsulfate (SDS) for tryptic digestion. The tryptic digestion of collection (I) was d(2)-^{13}C-formaldehyde labeled while the other 3 tryptic digests were d(0)-^{12}C-formaldehyde labeled.[26] The mass difference of d(2)-^{13}C-formaldehyde and d(0)-^{12}C-formaldehyde is 6.03. The d(2)-^{13}C-formaldehyde labeled tryptic digest from collection (I) was divided into 3 equal aliquots. Each aliquot was then mixed with one-third of the other 3 d(0)-^{12}C-formaldehyde labeled tryptic digests, respectively. Thus we obtained 3 peptide mixtures labeled with d(2)-^{13}C-formaldehyde and d(0)-^{12}C-formaldehyde. The 3 peptide mixtures were desalted with a Sep-Pac C-18 Cartridge from Waters Corporation (Milford, MA, USA) and then used for microbore reversed phase HPLC-MALDI QqTOF MS analysis, as described in section 4.2.4.

Figure 4.13 shows the distributions of the MS signal intensity ratios of the paired peptides that were labeled by d(2)-^{13}C-formaldehyde labeled and d(0)-^{12}C-formaldehyde in the three mixtures. Theoretically, all of the ratios should exactly be 1. Experimentally in Figure 4.13 the average ratio for the 3 mixtures is 1.01, 1.04, and 0.94, respectively, which agree with the theoretical value well and indicate the albumin depletion reproducibility with the extended pH-range chromatofocusing is satisfactory. Procedures like desalting, precipitating and dissolving proteins could bring about reproducibility errors and cause the ratio to deviate from 1.

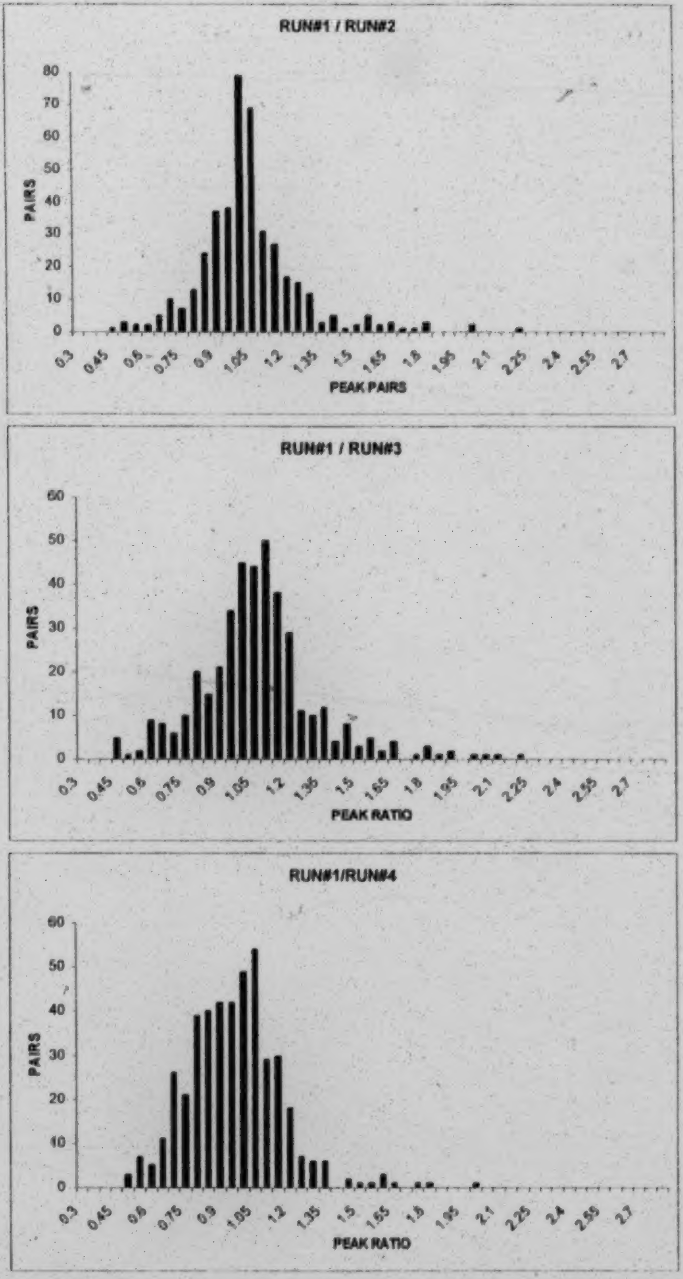


Figure 4.13. The distributions of the intensity ratios of the paired peptides that were d(2)-^{13}C-formaldehyde labeled and 3 d(0)-^{12}C-formaldehyde labeled in the 3 mixtures: 1/3 d(0)-^{12}C-formaldehyde labeled peptides from albumin depletion run #2, run #3, and run #4, respectively mixed with 1/3 d(2)-^{13}C-formaldehyde labeled peptides from albumin depletion run #1.

One MALDI plate was used for MALDI MS/MS and 296 peptides were identified through the isotope labeled pairs, of which 5 peptides were from Ig G1 and no peptide was found from albumin. The results indicate the albumin depletion efficiency with the extended pH-range chromatofocusing is high.

The number of peptides identified through the one-dimensional reversed phase HPLC-MALDI MS/MS analysis was only 296. This result shows that one dimensional separation strategy is not good enough for serum proteome analysis. The extended pH-range chromatofocusing method, acting as the first dimensional separation, was thus coupled with monolithic reversed phase HPLC-ESI MS/MS, which was discussed in Chapter 3, to establish a two-dimensional HPLC strategy for LC MS/MS analysis of human serum proteome.

Chromatofocusing fractions in a 4 min elution window were collected before and after the albumin-containing fraction (28 min to 39 min, as shown in Figure 4.11.). There were a total of 11 non-albumin fractions collected. Figure 4.14 shows the two approaches to further separate and identify the proteins in the 11 fractions by LC-ESI MSMS.

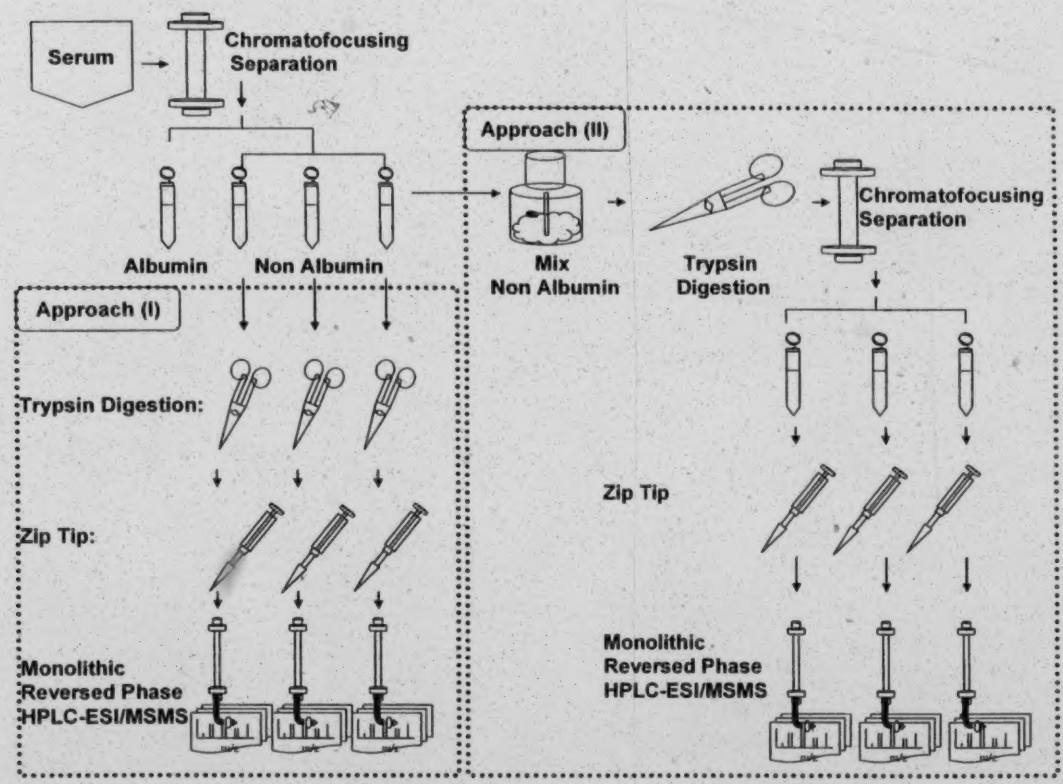

Serum
Chromatofocusing Separation
Albumin
Non Albumin
Approach (I)
Trypsin Digestion:
Zip Tip:
Monolithic Reversed Phase HPLC-ESI/MSMS
Approach (II)
Mix Non Albumin
Trypsin Digestion
Chromatofocusing Separation
Zip Tip
Monolithic Reversed Phase HPLC-ESI/MSMS

Figure 4.14. Schematic setup to separate and identify the proteins in the non-albumin fractions from the extended pH-range chromatofocusing by LC-ESI MS/MS on an LCQ Advantage.

Through approach (I), shown in Figure 4.14, 11 non-albumin fractions from the extended pH-range chromatofocusing were first digested by trypsin. The tryptic digests were then separated and identified through monolithic reversed phase LC-ESI MS/MS. With the 2-dimensional LC-ESI MSMS method, 2222 proteins were identified, of which >73% were exclusively identified from the combined protein identification data that were generated within one fraction only. The results illustrate that good separation of proteins was achieved.

In approach (II), the albumin-depleted protein mixture was trypsin digested and the tryptic digests were separated with the extended pH-range chromatofocusing on a 25 cm × 1 mm strong cation exchange column. The gradient was a 30 min A-B linear gradient. 22 chromatofocusing fractions at an interval of 1.5 min were collected and subjected to monolithic reversed phase HPLC-ESI MS/MS. With this 2-dimensional LC-ESI MS/MS method, 3783 peptides and 2570 proteins were identified, of which >58% were exclusively identified from the data generated in one fraction only, which illustrates that good separation of peptides was achieved.

The results of the two approaches are summarized in Table 4.6. Compared to approach (I), approach (II) uses the twice many fraction numbers, but the number of proteins identified is almost the same. After digestion the complexities of the peptide mixtures are expected to be at least 1 order of magnitude higher than the original protein mixtures, assuming that on average one protein is digested into 10 peptides. Although in approach (II) more fractions were collected for the reversed phase HPLC MS/MS, higher complexities in the peptide mixtures in approach (II) resulted in higher ion suppression in the reversed phase LC MS/MS, preventing more proteins from being identified.

Table 4.6. Number of proteins identified by approaches (I) and (II) shown in Figure 4.14.

Peptide Number Threshold	Number of Identified Proteins in Approach (I)	Number of Identified Proteins in Approach (II)	Number of Identified Proteins in total
>=1	2222	2570	3661
>=2	759	753	1269
>=3	447	484	729

4.4. Conclusions

A chromatofocusing method with a pH range of 3 to 12 has been developed. This pH range is wider than that achieved by the commercial chromatofocusing methods and covers almost all proteins' pIs. The pH linearity obtained by using different buffers on a strong cation exchange column was evaluated. The pH linearity was found to have a significant effect on chromatofocusing: better and flatter pH linearity leading to better chromatofocusing separation. The pH linearity simulation and the chromatofocusing mechanism were provided. Since the simulated pH curves have a good agreement with the experimental results, the model that the simulation was based on seems to be valid. This model and simulation can be used to predict the pH linearity as buffer compositions change, which is important for optimization of separation conditions and development of new buffer systems. The effects of column length on the chromatofocusing were studied. Longer HPLC columns could lead to better chromatofocusing separation since an extended focusing path could be obtained.

The performance of the extended pH-range chromatofocusing was compared with the ion exchange method on the same column. Due to the focusing effect, the chromatofocusing method's separation efficiency was higher than the regular ion exchange method. Because a buffer with low ionic strength was applied, the chromatofocusing method also has the following advantages: the regeneration time required for the column was far shorter than that for the regular ion exchange and the proteins would not be denatured during separation. The method can be implemented in conventional HPLC systems and no additional hardware is needed.

Standard proteins with a wide pI range were separated with the extended pH-range chromatofocusing method. A model was put forward to explain why the observed pIs were different from the theoretical pIs. The observed pIs were also related to the flow rate and the applied gradient. Slower gradient or flow rate made the observed pI closer to the theoretical pIs. The method can be potentially used to evaluate all proteins' experimental pIs since the method covers almost all protein pIs.

Through a one-dimensional separation strategy, the extended pH-range chromatofocusing method was used to separate milk components. A total of 139 proteins were detected by milk protein fractionation and MALDI-TOF MS analysis. This number is one order of magnitude higher than that obtained from direct analysis of the milk sample without the chromatofocusing separation.

This chromatofocusing method was found to be very efficient to deplete high abundance proteins such as albumin in serum. Through a two-dimensional separation strategy, either at the protein level separation (approach I) or at the peptide level separation (approach II), the extended pH-range chromatofocusing method was coupled with monolithic reversed phase LC-ESI MS/MS for human serum proteome analysis. Although in approach (II) more fractions were collected for the reversed phase HPLC MS/MS, higher complexity in the peptide mixtures in approach (II) resulted in higher ion suppression in MS/MS. In total, more than 3600 proteins were identified by the combination of the two approaches. Because the instrument used was an LCQ Advantage, a low-end MS/MS instrument, a better instrument, such as lqTOF, should lead to the identification of more proteins.

4.5. References

1. Yano, H. *Anal. Chem.* **2003,** 75, (17), 4682-4685.
2. Cooper, J. W.; Lee, C. S. *Anal. Chem.* **2004,** 76, (8), 2196-2202.
3. Li, Y.; Buch, J. S.; Rosenberger, F.; DeVoe, D. L.; Lee, C. S. *Anal. Chem.* **2004,** 76, (3), 742-748.
4. Bruchert, W.; Bettmer, J. *Anal. Chem.* **2005,** 77, (15), 5072-5075.
5. Chen, Y.; Kim, S. C.; Zhao, Y. *Anal. Chem.* **2005,** 77, (24), 8179-8184.
6. Chromy, B. A.; Gonzales, A. D.; Perkins, J.; Choi, M. W.; Corzett, M. H.; Chang, B. C.; Corzett, C. H.; McCutchen-Maloney, S. L., *J. Proteome Res.* **2004,** 3, (6), 1120-1127.
7. Yu, K. H.; Rustgi, A. K.; Blair, I. A., *J. Proteome Res.* **2005,** 4, (5), 1742-1751.
8. Sluyterman, L. A. A. E.; Wijdenes, J. *J. Chromatogr.* **1978,** 150, (1), 17-30.
9. Romijn, E. P.; Krijgsveld, J.; Heck, A. J. R. *J. Chromatogr. A* **2003,** 1000, (1), 589-608.
10. Shan, L.; Anderson, D. J. *Anal. Chem.* **2002,** 74, (21), 5641-5649.
11. Yan, F.; Subramanian, B.; Nakeff, A.; Barder, T. J.; Parus, S. J.; Lubman, D. M. *Anal. Chem.* **2003,** 75, (10), 2299-2308.
12. Shimohama, S.; Sasaki, Y.; Fujimoto, S.; Kamiya, S.; Taniguchi, T.; Takenawa, T.; Kimura, J. *Neuroscience* **1998,** 82, (4), 999-1007.
13. Berlin, Alejandro G.; Gusakov, Alexander V.; Sinitsyna, Olga A.; Sinitsyn, Arkady P. *Appl. Biochem. Biotechnol.* **2000,** 88, (1), 345-352.
14. Deng, Fan; Hatzios, Kriton K. *Pestic. Biochem. Physiol.* **2003,** 74, (2), 102-115.

15. Tabata, Toshiki; Katoh, Miki; Tokudome, Shogo; Nakajima, Miki; Yokoi, Tsuyoshi. *Drug Metabolism and Disposition* **2004,** 32, (10), 1103-1110.

16. Wang, Linsong; Burhenne, Kim; Kristensen, Brian K.; Rasmussen, Soren K. *Gene* **2004,** 343, (2), 323-335.

17. Pollard, Janet H. Scott; Katrina L. Kelner; Harvey B. *Anal. Biochem.* **1985,** 149, 163-165.

18. Liu, Y.; Anderson, D. J. *J. Chromatogr. A* **1997,** 762, (1), 47-54.

19. Strong, J. C.; Frey, D. D. *J. Chromatogr. A* **1997,** 769, (2), 129-143.

20. Zhang, B.; McDonald, C.; Li, L. *Anal. Chem.* **2004,** 76, 992-1001.

21. Kreunin, P; Urquidi, V.; Lubman, D. M.; Goodison, S. *Proteomics* **2004,** 4, 2754–2765.

22. Gerard J. P.; Bertoglio J., *J Immunol Methods.* **1982,** 55, (2), 243-251.

23. Hooks, M. A.; Bode, K.; COUÉE, I. *Biochem. J.* **1996,** 320, 607-614.

24. Zhu, K.; Zhao, J.; Lubman, D. M.; Miller, F. R.; Barder, T. J. *Anal. Chem.* **2005,** 77, (9), 2745-2755.

25. Anderson, N. L.; Anderson, N. G. *Mol. Cell. Proteomics* **2002,** 1, 845-867.

26. Ji, C.; Guo, N.; Li, L. *J. Proteome Res.* **2005,** 4, 2099-2108.

Chapter 5

Conclusions and Future Work

Multidimensional separation has been studied for many years since its introduction in the form of paper chromatography. Today multidimensional separation techniques, such as gel electrophoresis, continuous elution tube gel electrophoresis/reversed phase HPLC, capillary zone electrophoresis/gel electrophoresis and microfluidic devices, have been successfully demonstrated in a variety of applications. However, the most mature and widely used multidimensional techniques for small biomolecules are based on HPLC systems.

Multidimensional separation provides a strategy to separate analytes that are too close in physical or chemical characteristics to be separated by a one-dimensional separation. HPLC using a bifunctional column has been demonstrated to provide an alternative technique for multidimensional separation of complex mixtures. The work in Chapter 2 shows that a bifunctional HPLC column can provide both single-mode and mixed-mode separation mechanisms. Method development in using the bifunctional HPLC column should start with optimization of single-mode separation, instead of mixed-mode separation, unless a single-mode separation mechanism cannot fulfill the objective, because the complexity of mixed-mode HPLC method development is much higher than that of a single-mode HPLC method.

Very polar small molecules, such as amino acids, have very poor retention on a reversed phase column. Ion-pairing reagents can be used to form ion-pairs with these polar compounds so that they can have usable retention times on a reversed phase HPLC column. Nevertheless the work in Chapter 2 suggests that in cation exchange-reversed phase mixed-mode HPLC the ion-pair reagents can increase retention of polar molecules only when the ion-pair reagents are at a concentration greater than their critical concentration. If the concentration of an ion-pair reagent is below its critical concentration, the ion-pair reagent may actually reduce the analytes' retention on the reversed phase HPLC column. In Chapter 2 it is also demonstrated that the retention enhancement caused by the ion-pairing reagents is proportionally related to the pI of the analytes, the concentration and length of the ion-pairing reagent. In the future a further study should be performed to confirm critical concentrations for the ion-pairing reagents in reversed phase HPLC and to build a more complex quantitative model so that the effects of ion-pairing reagents on the retention of analytes on a mixed-mode HPLC column can be quantitatively predicted and the method development for separating some specific target compounds may be greatly accelerated.

In Chapter 2 multidimensional HPLC separation using a bifunctional column combined with MS detection is described. Compared with regular on-line multidimensional HPLC, this multidimensional HPLC system using a single HPLC column should lower costs significantly by eliminating the need for multiplex HPLC systems, columns and on-line traps. A post-column TFA-fix is used to enhance the method's sensitivity for amino acid analysis. Different TFA-fixes are evaluated and a mechanism is described to explain their performance. High efficient separation of 20 amino acids and some metabolites is demonstrated. Identification and quantification of these molecules in standard mixtures and human urine samples are done by using ESI

MS. In the future a better MS/MS detection mode with multiple reactions monitoring (MRM) should be used to replace the MS detection mode to improve sensitivity and specificity for amino acid analysis.

Chapter 3 describes a new method based on the use of monolithic column HPLC and ESI MS/MS for fast bacterial identification. Rapid detection and identification of bacteria plays a crucial role in medical microbiology and detection of biohazards. Both spectroscopy and MS have been extensively studied in microorganism detection through protein identification. However, only MS has the ability of bacterial spore identification and fingerprinting identification. In the method described in Chapter 3, fast and efficient separation of the peptides from the digestion of bacterial extracts is achieved through reversed phase HPLC with a capillary monolithic column. Benefiting from the unique biporous structure, the monolithic column is demonstrated to have very high separation efficiency and can be used to separate tryptic digests of bacterial proteins with gradients as fast as 2.5 min. It is demonstrated that, in a Finnigan ion trap mass spectrometer, a combination of 2 microscans for each MS survey scan, followed by 4 analytical MS/MS scans is the best compromise between data acquisition speed and spectral quality. The total required analysis time is much shorter than when conventional particle packed columns are used. This method is used for identifying pure bacterial cultures and simple mixtures of bacteria.

In Chapter 4 a chromatofocusing method with a pH range of 3 to 12 is described. This pH range is wider than the commercial chromatofocusing methods and coveres almost all proteins' pIs. The pH linearity, achieved by using different buffers on a strong cation exchange column, is evaluated and found to have a significant effect on chromatofocusing efficiency. A more linear pH change during the gradient experiment provides better chromatofocusing separation. A separation model and a computer simulation method of the developed chromatofocusing method are discussed. The simulation results match well with the experimental data. Thus, the simulation method can be used to optimize separation conditions and develop new buffer systems.

In Chapter 4, a comparison of the performance of the extended pH-range chromatofocusing with an ion exchange method on the same column is described. Due to the focusing effect, the chromatofocusing method's separation efficiency is shown to be superior to the regular ion exchange method. Because a buffer with low ionic strength is applied, the chromatofocusing method also has several other advantages. The regeneration time required for the column is far shorter than that for regular ion exchange chromatography. The proteins would not be readily denatured during separation. This extended pH-range chromatofocusing method used a buffer system that was also carefully developed so that its species would not bind or react with analytes, and would not interfere with any next dimension of separation. In contrast, the commercial chromatofocusing system, that uses Polybuffers, can lead to complications from removing these Polybuffer components from the eluate after chromatofocusing by requiring a separate dimension of HPLC to remove the Polybuffer components.

The results shown in Chapter 4 also show that the extended pH-range chromatofocusing method is very efficient at depleting abundant proteins, such as albumin, in human serum. Using a two-dimensional separation strategy, either at the protein level separation (approach I) or at the peptide level separation (approach II), the extended pH-range chromatofocusing method is coupled with monolithic reversed phase

HPLC-ESI MS/MS for human serum proteome analysis. A total of more than 3600 proteins have been identified by using the combination of approachs (I) and (II). For future work, we expect that an even greater number of proteins can be identified by using a better MS/MS instrument, such as Q-TOF mass spectrometer, rather than an LCQ Advantage used in this work.